FORM PTO-1390 (REV. 9-2001) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: Mo-6878/LeA 33,759

U.S. APPLICATION NO. (If known, see 37 CFR 1.5): 10/030566 To Be Assigned

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/EP00/06104	30 June 2000 (30.06.00)	09 July 1999 (9.07.99)

TITLE OF INVENTION: DNA CODING FOR -TUBULIN AND USE THEREOF

APPLICANT(S) FOR DO/EO/US: SAMSON-HIMMELSTJERNA, Georg von; HARDER, Achim; SCHNIEDER, Thomas and PAPE, Michaela

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. [X] This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. [X] The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. [X] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [X] is attached hereto (required only if not communicated by the International Bureau).
   b. [ ] has been communicated by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [X] An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. [X] is attached hereto.
   b. [ ] has been previously submitted under 35 U.S.C. 154(d)(4).
7. [ ] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. [ ] are attached hereto (required only if not communicated by the International Bureau).
   b. [ ] have been communicated by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. [X] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. [ ] An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 20 below concern document(s) or information included:

11. [ ] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [X] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [ ] A FIRST preliminary amendment.
14. [ ] A SECOND or SUBSEQUENT preliminary amendment.
15. [ ] A substitute specification.
16. [ ] A change of power of attorney and/or address letter.
17. [ ] A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.
18. [ ] A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. [ ] A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. [X] Other items or information:

Statement of Verification w/paper and disk copy of sequence listing

U S APPLICATION NO. (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
To Be Assigned	PCT/EP00/06104	Mo-6878/LeA 33,759

21. [X] The following fees are submitted:	CALCULATIONS PTO USE ONLY
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)):	
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1040.00**	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$890.00**	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO **$740.00**	
International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$710.00**	
International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**	
ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 890.00
Surcharge of **$130.00** for furnishing the oath or declaration later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$ 0.00

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	45 - 20 =	25	x **$18.00**	$ 450.00
Independent claims	2 - 3 =	0	x **$84.00**	$ 0.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$280.00**	$ 280.00
TOTAL OF ABOVE CALCULATIONS =				$ 1,620.00
[ ] Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.			+	$ 0.00
SUBTOTAL =				$ 1,620.00
Processing fee of **$130.00** for furnishing the English translation later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				$ 0.00
TOTAL NATIONAL FEE =				$ 1,620.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$ 40.00
TOTAL FEES ENCLOSED =				$ 1,660.00
			Amount to be refunded:	$
			charged:	$

a. [ ] A check in the amount of $ _____ to cover the above fees is enclosed.

b. [X] Please charge my Deposit Account No. 13-3848 in the amount of $ 1,660.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [X] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 13-3848 . A duplicate copy of this sheet is enclosed.

d. [ ] Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

00157
PATENT TRADEMARK OFFICE

SIGNATURE

Godfried R. Akorli
NAME

28,779
REGISTRATION NUMBER

PATENT APPLICATION
Mo-6878
LeA 33,759

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION OF	)
	) PCT/EP00/06104
GEORG VON SAMSON-HIMMELSTJERNA ET AL	)
	)
SERIAL NUMBER: TO BE ASSIGNED	)
	)
FILED: HEREWITH	)
	)
TITLE: DNA CODING β-TUBULIN AND USE THEREOF	)

STATEMENT OF VERIFICATION OF SEQUENCE LISTING

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

Applicants submit this statement as verification that the simultaneously submitted paper copy and disk copy of the subject sequence listing are identical.

Respectfully submitted,

By _____
Godfried R. Akorli
Attorney for Applicants
Reg. No. 28,779

Bayer Corporation
100 Bayer Road
Pittsburgh, Pennsylvania 15205-9741
(412) 777-8340
FACSIMILE PHONE NUMBER:
(412) 777-8363
s:\ksl\6878verification

"Express Mail" mailing label number EV700175964US

Date of Deposit January 7, 2002

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents and Trademarks, Washington, D.C. 20231

Donna J. Veatch
(Name of person mailing paper or fee)

_____
Signature of person mailing paper or fee)

"Express Mail" mailing label number EL700175964US
Date of Deposit January 7, 2002

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents and Trademarks, Washington, D.C. 20231

Donna J. Veatch
(Name of person mailing paper or fee)

Signature of person mailing paper or fee)

10/030566

JC13 Rec'd PCT/PTO 07 JAN 2002

DNA coding for ß-tubulin and its use

The invention relates to DNA which codes for β-tubulin from nematodes of the family of the Strongylidae, the polypeptide encoded by this DNA, the use of the DNA for the diagnosis of anthelmintic resistance of these nematodes and for the identification of the species of these nematodes, the use of the β-tubulin as a constituent of a vaccine, and a process for the identification of new anthelmintic or antibiotic compounds.

Parasitic helminths (worms) are a health problem for humans and animals and cause significant economic damage. In 1996, the expenditure for anthelmintics worldwide was more than 2 billion US dollars. The most important anthelmintics which are presently used can be divided into three groups according to their mechanism of action:

1. The cyclic amidines pyrantel and morantel act, together with the imidazothiazoles tetramisole and levamisole, as cholinergic compounds for the parasitic nervous system.

2. The benzimidazoles are inhibitors of the polymerization of microtubules and lead to the degradation of tubulin, followed by the loss of a number of cell functions such as transport within cells and cell division.

3. The macrocyclic lactones bind and open glutaminergic chloride channels and thus act as inhibitors of the nervous system of nematodes and arthropods.

Microtubules are built up from tubulin subunits. Tubulin is a dimeric protein which consists of α- and ß-tubulin and is in a dynamic equilibrium between tubulin and microtubules. This equilibrium can be influenced by exogenous substances, which are designated as microtubule inhibitors. Some of these inhibitors, such as, for example,

the benzimidazoles, act by their binding to tubulin, whereby they prevent the self-association of these subunits to growing microtubules, while at the opposite end the dissociation of the microtubules is continued. On account of this, dysfunctions occur in processes within the cell which are important to life and finally the death of the cell and of the entire organism occurs (Lacey, E. (1990) Mode of action of benzimidazoles. Parasitology Today 6, 112-115). Such microtubule inhibitors include various classes of compound which are synthetically prepared or produced by various organisms.

The binding of microtubule inhibitors to tubulin from various organisms shows great differences with respect to the affinity of the binding. Thus the anthelmintics oxfendazole and thiabendazole show a high affinity for tubulin from *Ascaridia galli* and an only slight affinity for tubulin from mammals such as the sheep (Dawson et al. (1983) Purification and characterisation of tubulin from the parasitic nematode, *Ascaridia galli*, Molecular and Biochemical Parasitology 7, 267-277). The selective toxicity of benzimidazoles can be explained by means of this selective affinity (Lacey, E. (1988) The role of the cytoskeletal protein, tubulin, in the mode of action and mechanism of drug resistance to benzimidazoles, International Journal for Parasitology 18, 885-936).

The widespread use of these anthelmintics has led to considerable resistance problems against all three classes, especially in livestock (Bauer et al. (1994) Anthelmintic resistance in nematodes of farm animals. A seminar organised for the European Commission, Brussels, Belgium, 8 to 9 November 1993, pp. 17-24). The most widespread class of anthelmintics are the benzimidazoles. Resistance to benzimidazoles has been described worldwide in the case of parasites of sheep, cattle, pigs and horses. Benzimidazoles are broad-spectrum anthelmintics with action against nematodes, cestodes and trematodes. Investigations in German stud farms showed a resistance of small strongylids to benzimidazoles in more than 80% of the cases (Ullrich et al. (1988) Benzimidazole resistance in small strongylids

(Cyathostominae): distribution in horse stock in Northrhine-Westphalia, Berliner Münchner Tierärztliche Wochenschrift 101, 406-408).

For an effective treatment with anthelmintics, it is therefore of great importance to obtain information about possible resistance of worms in a horse or within a herd of horses. In order to check the possible resistance of a worm population of small stronglids, diagnostic procedures have already been developed which are based on the activity of anthelmintics in the stages of development of the parasites (eggs, larvae) (Coles et al. (1992) World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.) Methods for the detection of anthelmintic resistance in nematodes of veterinary importance, Veterinary Parasitology, 44, 35-44). The "larval development assay" (LDA) describes the inhibitory effect of anthelmintics in the nonparasitic first larval stage as a function of the anthelmintic concentration employed. In a similar manner, the second approach, the "egg hatch assay" (EHA) uses the inhibitory action of the anthelmintics on the hatching of the larvae as a function of the anthelmintic concentration employed. A disadvantage of both approaches lies in the low reproducibility of the results. Moreover, both approaches are time-consuming and labor-intensive.

The sensitivity of both approaches is moreover low. An existing resistance is only detected if more than 25% of the population are resistant (Roos et al. (1995) New genetic and practical implications of selection for anthelmintic resistance in parasitic nematodes, Parasitology Today 11, 148-150).

There is therefore an urgent need for diagnostic procedures which are sensitive, rapid and reproducible. For the sheep parasites, *Haemonchus contortus* and *Teladorsagia circumcincta*, a procedure based on the PCR technique has been described. This is based on at least one point mutation in codon 200 of the ß-tubulin isotype 1 gene (Elard et al. (1999) PCR diagnosis of benzimidazole - susceptibility or - resistance in natural populations of the small ruminant parasite, *Teladorsagia circumcincta*, Veterinary Parasitology 80, 231-237). The point mutations in codon 200 result in an

amino acid replacement of phenylalanine by tyrosine and correlates with a benzimidazole resistance of the mutated protein (Kwa et al. (1995) ß-tubulin genes from parasitic nematode *Haemonchus contortus* modulate drug resistance in *Caenorhabditis elegans*, Journal of Molecular Biology 246, 500-510).

From studies on *Haemonchus contortus,* it is known that the sequences of nematodes coding for tubulin are species-specific (WO 92/03549).

The various species of the small strongylids of the horse differ in their epidemiological frequency. The species which are most important and found most frequently worldwide include *Cylicocyclus nassatus*, *Cyathostomum coronatum* and *Cyathostomum catinatum*. Resistances of these species to benzimidazoles have been described in various countries, inter alia also in Germany (Burger, H.-J. and Bauer, C. (1987) Efficacy of four anthelmintics against benzimidazole - resistant cyathostomes of horses, Veterinary Record 120, 293-296).

The nucleic acid sequencing of ß-tubulin cDNAs from the species of small strongylids according to the invention had an identity of over 95%. The identity with the known, abovementioned ß-tubulin sequences of sheep parasites, however, is only 75.4-82.6%.

The derived amino acid sequences are very similar within the sequences according to the invention. This is also true of the derived ß-tubulin amino acid sequences of the sheep parasites. The identity here is between 95 and 99.8%. Only very few positions result in which amino acid exchange occurs. To be emphasized here is codon 200, in which a change from phenylalanine to tyrosine results.

However, the noncoding ß-tubulin sequences from various helminth species published hitherto show no significant identity. It is accordingly surprising that not only the coding sequences, but also parts of the noncoding sequences, of the various species of small strongylids of the application present here have a high identity.

These regions are therefore also suitable to be able to differentiate various species of small strongylids and other nematode species from one another. PCR primers which are derived from these intron regions can be used for the specific detection of small strongylids within a sample which also contains genetic material from other helminth organisms.

ß-Tubulin from nematodes or parts thereof are known for having a protective, immunological potential (Bughio et al. (1993) Characterisation and biological activities of anti-*Brugia pahangi* tubulin monoclonal antibodies, International Journal for Parasitology, 7, 913-924). The ß-tubulin of the small strongylids encoded by the abovementioned DNA can be used as a vaccine just as monoclonal antibodies can be used against the ß-tubulin.

Inhibitors of the interaction of the tubulin or of its subunits, such as benzimidazoles, colchizine and taxol, are important lead structures of a series of therapeutics which are directed against human, animal or plant diseases. The importance of tubulin as the target of these compounds is far-reaching and underlines its potential for the search for new active compounds for the control of these diseases.

The invention relates to the DNA coding for the ß-tubulin from nematodes of the family of the Strongylidae, particularly of the subfamily of the Cyathostominae, or fragments of this DNA. The ß-tubulin DNA can in this case be genomic DNA or cDNA. The DNA sequences which this invention relates to can be considered as new members of the tubulin gene family of parasitic nematodes of the order Strongylida, particularly of the subfamily of the Cyathostominae. The invention relates very particularly to the DNA sequences which code for ß-tubulin from parasitic nematodes of the genera Cyathostomum and Cylicocyclus.

The invention likewise relates to DNA sequences which have an identity of more than 85% to a polynucleotide coding for one of the amino acid sequences as set forth in SEQ ID NO. 2, 4, 6, 8 or 10.

The invention likewise relates to preferred DNA sequences which have an identity of more than 95% to a polynucleotide coding for one of the amino acid sequences as set forth in SEQ ID NO. 2, 4, 6, 8 or 10.

The invention relates in particular to DNA sequences coding for ß-tubulin, which originate from parasitic nematodes of the genera Cylicocyclus and Cyathostomum, very particularly those sequences which originate from parasitic nematodes of the species *Cylicocyclus nassatus*, preferably DNA as set forth in SEQ ID NO. 3, 5, 7, 9 or 11 or from *Cyathostonum coronatum*, preferably DNA according to SEQ ID NO. 1.

The invention likewise relates to DNA sequences as described above, which in contrast to these sequences have at least one point mutation or one nucleotide replacement in codon 200. These point mutations result in a change in the amino acid sequence encoded by this DNA, e.g. a replacement of the amino acid phenylalanine by tyrosine, and correlate with the resistance of tubulin with appropriate mutations to benzimidazoles.

The invention likewise relates to DNA sequences which are complementary to the DNA described above or fragments of this DNA, and to fragments of these DNA sequences. These DNA sequences or these fragments comprise oligonucleotides which are derived from one of the DNA sequences mentioned above or described under SEQ ID NO. 1, 3, 5, 7, 9 or 11 or sequences identical to 85% thereto, preferably sequences identical to 95%, and are derived from strands complementary thereto and can hybridize to these.

The invention in this case relates preferably to oligonucleotides consisting of or comprising one of the sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51, which hybridize to abovementioned DNA sequences, preferably in the region of noncoding sequence sections of the ß-tubulin genes.

The invention likewise preferably relates to oligonucleotides consisting of or comprising one of the sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51, which hybridize to coding regions of the abovementioned sequences.

The invention likewise relates to RNA sequences which are complementary to the DNA described above or fragments of this DNA, and fragments of these RNA sequences. These RNA sequences or these fragments comprise ribooligonucleotides which correspond to a region of one of the DNA sequences mentioned above or described under SEQ ID NO. 1, 3, 5, 7, 9 or 11, sequences complementary thereto or DNA sequences 85% identical, preferably 95% identical, thereto, and can hybridize to these.

The invention likewise relates to an expression construct which comprises one of the DNA sequences described above, and to a DNA sequence linked thereto which makes possible the expression of the DNA. These include, for example, at least one promoter for constitutive or inducible expression or alternatively enhancers. Suitable promoters for expression in *E. coli* are natural hybrid or bacteriophage promoters, preferably promoters of the group of λ-phages, hsp, omp or synthetic promoters such as mentioned, for example, in WO 98/5625, DE 3 430 683 or EP 0 173 149.

The invention likewise relates to vectors which comprise one of the DNA sequences described above and make possible the expression of the ß-tubulin according to the invention or fragments thereof in a host cell.

The invention likewise relates to host cells which contain the abovementioned DNA, an expression construct as mentioned above, or a vector and allow the expression of the ß-tubulin or fragments thereof.

The invention likewise relates to polypeptides which are encoded by one of the abovementioned DNA sequences or fragments of these DNA sequences, and to fragments of these polypeptides.

The invention in this case relates preferably to polypeptides which are encoded by a DNA sequence comprising SEQ ID NO. 1, 3, 5, 7, 9 or 11, by DNA sequences which have an identity of 85% to these sequences, preferably of 95%, or of fragments of this DNA.

The invention also relates to polypeptides which are encoded by the DNA sequence described above, which contain at least one point mutation in codon 200 as described above and show resistance to benzimidazoles, and to fragments of these polypeptides.

The invention in this case relates very particularly preferably to polypeptides comprising one of the amino acid sequences described in SEQ ID NO. 2, 4, 6, 8 or 10 or fragments thereof.

The invention in this case relates to polypeptides, particularly to purified polypeptides or polypeptides prepared recombinantly.

The invention relates to polypeptides of full length and also to corresponding fragments of these polypeptides, e.g. certain motifs or domains. These fragments can be of different length and comprise, for example, 5, 10, 25, 50, 100, 150, 200, 250 or 300 amino acids.

This invention likewise relates to fusion proteins which comprise a polypeptide as described above. The fusion protein can in this case contain a further polypeptide component which is not connected to the ß-tubulin (e.g. LexA, B42, glutathione S-transferase, a His tag, a polypeptide having enzymatic activity such as alkaline phosphatase or an epitope tag).

The invention also relates to a process for the preparation of a polypeptide as described above in suitable prokaryotic or eukaryotic expression systems. The expression can in this case be carried out permanently or transiently as described above in a corresponding cell line in each case or corresponding host cells. Suitable prokaryotic expression systems are known host-vector systems such as bacteria (e.g. *Streptomyces spp., Bacillus subtilis, Salmonella typhimurium, Serratia marcescens* and particularly *Escherichia coli*).

Expression in a eukaryotic system is preferably carried out in the baculovirus system, particularly in a system which allows the introduction of post-translational modifications.

This invention likewise relates to the use of DNA as mentioned above for the detection of DNA from nematodes of the family Strongylidae, preferably of the subfamily Cyathostominae, particularly preferably of the genera Cyathostomum and Cylicocyclus, very particularly preferably of the species *Cyathostomum coronatum* and *Cylicocyclus nassatus*. The invention in this case relates to oligonucleotides as mentioned above which are complementary to DNA coding for ß-tubulin or strands complementary thereto and can hybridize to this DNA. Preferably, these oligonucleotides hybridize to the intron regions, i.e. the noncoding DNA sequences. The invention relates to the use of these oligonucleotides or parts thereof as

a) samples in Northern or Southern blot assays,

b) PCR primers in a diagnostic procedure for the detection of the abovementioned nematodes, the DNA of the nematodes concerned being specifically identified and amplified with the aid of the primer and the PCR technique.

The invention in this case relates preferably to oligonucleotides consisting of or comprising one of the sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51.

The invention likewise relates to the use of DNA as mentioned above for the detection of DNA from nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, which codes for ß-tubulin or fragments thereof, which are resistant to benzimidazoles. The invention in this case relates to oligonucleotides as mentioned above which are complementary to DNAs which code for ß-tubulin with a resistance to benzimidazoles or to the complementary strands of this DNA and which can specifically hybridize to this DNA.

The invention also relates to the use of these oligonucleotides or parts thereof as

a) samples in Northern or Southern blot assays,

b) PCR primers in a diagnostic procedure for the detection of the abovementioned nematodes with a resistance to benzimidazoles, the DNA of the nematodes concerned being specifically identified and amplified with the aid of the primer and the PCR technique.

The invention in this case preferably relates to oligonucleotides consisting of or comprising one of the sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51.

The invention also relates to a procedure for the detection of nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, where oligonucleotides as described above specifically hybridize to DNA sequences which originate from the organisms mentioned, and which can be amplified with the aid of

the PCR technique. The hybridization is preferably carried out in the noncoding regions of the β-tubulin gene (introns).

The detection of organisms as mentioned above can be carried out, for example, by

a) making available an oligonucleotide probe or primer which can hybridize to the abovementioned DNA coding for ß-tubulin or strands complementary thereto or to the 5'- or 3'-flanking regions thereof,

b) bringing the oligonucleotide probe or the primers into contact with an appropriately prepared probe containing DNA,

c) detecting the hybridization of the oligonucleotide or primer (e.g. with the aid of the polymerase chain reaction),

d) sequencing the detected sequence of the ß-tubulin gene, and

e) comparing this sequence with the DNA sequences according to the invention which were described above, preferably with DNA sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11.

The invention also relates to a process for the detection of nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, which are resistant to benzimidazoles, where oligonucleotides as described above specifically hybridize to DNA sequences which originate from the organisms mentioned and which can be amplified with the aid of the PCR technique. The hybridization is preferably carried out in the noncoding regions of the β-tubulin gene (introns).

The detection of organisms as mentioned above can be carried out, for example, by

a) making available an oligonucleotide probe or primer which can hybridize to the abovementioned DNA coding for ß-tubulin or strands complementary thereto or to the 5'- or 3'-flanking regions thereof,

b) bringing the oligonucleotide probe or the primers into contact with an appropriately prepared probe containing DNA,

c) detecting the hybridization of the oligonucleotide or primer (e.g. with the aid of the polymerase chain reaction),

d) sequencing the detected sequence of the ß-tubulin gene, and

e) comparing this sequence with the DNA sequences according to the invention which were described above, preferably with DNA sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, having at least one point mutation in codon 200, which leads to a resistance of the β-tubulin encoded by these sequences to benzimidazoles.

The invention likewise relates to a diagnostic test kit for the detection and identification of nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, which, inter alia, makes available oligonucleotides as described above, which can be used in procedures for the detection of the species mentioned. The present invention likewise makes available oligonucleotides which specifically hybridize to sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, sequences complementary thereto, sequences having at least one point mutation in codon 200, or fragments of these sequences. In this case, oligonucleotides consisting

of or including sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51 are particularly preferred.

The invention likewise relates to a diagnostic test kit for the detection of nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, having a resistance to benzimidazoles which, inter alia, makes available oligonucleotides as described above, which can be used in procedures for the detection of the species mentioned. The present invention likewise makes available oligonucleotides which hybridize specifically to sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, sequences complementary thereto, sequences having at least one point mutation in codon 200, or fragments of these sequences. In this case, oligonucleotides consisting of or comprising sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51 are particularly preferred.

The invention likewise relates to a diagnostic test kit as described above, where the oligonucleotides made available in this kit are provided with a detectable marker. Such detectable markers can include, inter alia, enzymes, enzyme substrates, coenzymes, enzyme inhibitors, fluorescence markers, chromophores, luminescent markers and radioisotopes.

This invention likewise relates to antibodies which react specifically with an epitope of a ß-tubulin from nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*.

This invention likewise particularly relates to monoclonal antibodies which react specifically with an epitope of a ß-tubulin from nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly

preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*.

This invention likewise relates to the use of the abovementioned antibodies as nematicides.

This invention also relates to the use of the abovementioned ß-tubulin polypeptides or fragments thereof from nematodes of the family of the Strongylidae, preferably of the subfamily of the Cyathostominae, particularly preferably of the genera Cylicocyclus and Cyathostomum, very particularly preferably of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, for the preparation of vaccines which contain at least one of the ß-tubulin polypeptides mentioned or fragments thereof. The vaccine is in this case able to produce an immune response which is specific for a ß-tubulin described above.

In a preferred embodiment, the vaccine contains an antigenic determinant, e.g. an individual determinant of a polypeptide having an amino acid sequence as set forth in SEQ ID NO. 2, 4, 6, 8 or 10 or of a polypeptide which is encoded by one of the abovementioned DNAs or fragments thereof.

The invention likewise relates to a procedure for the preparation of an immunogenic composition for the immunization of mammals, consisting of at least one of the abovementioned ß-tubulin polypeptides according to the invention or fragments thereof or of at least one of the abovementioned antibodies.

The invention likewise relates to the use of the expression vectors described above containing a nucleic acid coding for a β-tubulin according to the invention, preferably a sequence as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, fragments thereof or sequences homologous thereto for the preparation of an immunogenic composition for administration to a host for the activation of a protective immune response in this host, which is directed at β-tubulin from parasitic nematodes.

The invention likewise relates to an immunogenic composition comprising a vector (comprising a nucleic acid coding for the β-tubulin according to the invention, preferably a sequence as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, fragments thereof or sequences homologous thereto and to a promoter sequence which is linked functionally to said nucleotide sequence and which controls the expression of the β-tubulin according to the invention producing an immune response) and a vehicle suitable for pharmaceutical purposes.

This invention likewise relates to a process for the identification of substances which modulate the interaction of tubulin or the interaction of subunits of tubulin. The process is based on the use of tubulin, preferably on tubulin from parasitic nematodes, particularly preferably on ß-tubulin from parasitic nematodes of the order strongylida, very particularly preferably on ß-tubulin from parasitic nematodes of the family of the Strongylidae, most preferably on ß-tubulin from parasitic nematodes of the subfamily of the Cyathostominae. A particularly preferred group of the ß-tubulin used in this procedure is ß-tubulin from parasitic nematodes of the genera Cylicocyclus and Cyathostomum.

The invention relates to the identification of substances, e.g. small organic molecules, which are able to modulate the interaction of tubulin protein molecules or its subunits with one another. Preferably, the invention relates to the identification of compounds which inhibit the interaction.

The invention likewise relates to a procedure as described above, which is based on

a) bringing the substance to be tested into contact with the tubulin, the conditions selected allowing the interaction of the tubulin molecules with one another and the binding of the test substance to tubulin,

b) detecting the binding which has taken place by determining the ability of the tubulin protein molecules to interact with one another and

c) comparing the ability of the tubulin protein molecules to interact with one another in the presence of a test substance to the ability to interact with one another in the absence of a test substance.

The invention likewise relates to a process for the identification of substances which modulate the ability of tubulin molecules to interact with one another. Particularly preferably, the invention in this case relates to a procedure which uses one of the polypeptides described above which are coded by the DNAs described above or fragments thereof, particularly by DNAs consisting of or comprising sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, and by sequences which have an identity of 85% thereto, preferably of 95% and code for ß-tubulin which has an amino acid sequence as set forth in SEQ 2, 4, 6, 8 or 10.

The invention likewise relates to a procedure for the identification of substances which modulate the ability of tubulin to interact with one another as described above, the procedure used being based on detecting a modulation of the tubulin interaction in the presence of a test substance with the aid of a test system based on cells. A preferred embodiment of such a test system is the "two hybrid system" (US 5 283 317, Zervos et al. (1993) Cell 72, 223-232; WO 94/10300). This system is suitable for documenting or describing the interaction of two proteins by the interaction leading to a detectable signal. Such a system can also be adapted to test systems having high throughput numbers.

The invention likewise relates to a procedure for the identification of substances which modulate the ability of tubulin to interact with one another, the procedure used being based on detecting a modulation of the tubulin interaction in the presence of a test substance with the aid of a cell-free test system. A particularly preferred embodiment of such a test system is the "scintillation proximity assay" (SPA)

(EP 015 473). This test system is based on the detection of an interaction of a receptor bound to microspheres or beads, e.g. of a tubulin molecule with a ligand, the microspheres or beads being provided with a scintillating molecule. A signal is detected if the receptor-ligand complex decomposes.

The invention likewise relates to substances which have hitherto still not been described, which are identified with the aid of the procedure described above and are suitable to modulate, preferably to inhibit, the interaction of tubulin molecules.

The invention likewise relates to the use of substances which have hitherto still not been described, which were identified using one of the procedures described above, for the preparation of an agent which is used for the prophylactic or therapeutic treatment of animals or humans which can be attacked or have been attacked by nematodes. The agents according to the invention contain at least one of the substances identified by one of the procedures described above and can be administered nasally, dermally, parenterally or enterally.

For better understanding, the meaning of certain words and terms which are used in the description, the examples and attached claims are explained in greater detail below.

The term "fragments" with respect to proteins and DNA describes parts of the nucleic acids or amino acid sequences described under SEQ ID NO. 1 to 11, sequences complementary thereto or sequences identical thereto to 85%, preferably to 95%. The fragments of the DNA and polypeptide sequences comprise at least five nucleotides or amino acids, but can likewise comprise up to 447 amino acids or up to 1343 nucleotides, or up to 2565 nucleotides in the case of the sequence as set forth in SEQ ID NO. 11.

The terms "homology", "identity" or "similarity" relate to sequence similarities between two peptides or between two nucleic acid molecules. Homology can be

determined by in each case comparing a position in each sequence with one another. If a position in the compared sequence is occupied by the same base or amino acid, the two molecules in this position are homologous. The measure of homology between sequences is a function of the number of the corresponding or homologous positions which the sequences share with one another. A "nonhomologous" sequence has an identity of less than 40%, but preferably of less than 25% identity.

The term "homology" in particular means that DNA segments of a length of at least 15 base pairs or strands complementary to the DNA agree to at least 85%, preferably 95%, with the nucleotides with the corresponding DNA. A homology can be determined, inter alia, with the aid of computer programs such as the GCG program (Devereux et al. (1983), Nucleic Acids Res. 12, 387-395).

A "homology" also exists if a DNA segment can hybridize to the DNA strand concerned or its complementary strand.

The term "hybridize" or "hybridization" describes the process in which a single-stranded nucleic acid molecule undergoes base pairing with a complementary DNA strand, the ability of a single-stranded nucleic acid molecule depending on the stringency of the hybridization conditions.

The term "stringency" relates to the hybridization conditions. "High stringency" is imparted if base pairing is rendered difficult. "Low stringency" is imparted if base pairing is facilitated.

The term "complementary" relates to the ability of purine and pyrimidine nucleotides to form base pairs with one another via hydrogen bridges. Complementary base pairs are, inter alia, guanine and cytosine, adenine and thymine and adenine and uracil.

The person skilled in the art realizes that, on account of the degenerate genetic code (i.e. 64 codons code for 20 amino acids), numerous "silent" substitutions of

nucleotide base pairs can be introduced into the sequence coded therefor without changing the identity of the protein products encoded thereby. All such substitutions are to be contained in the scope of the invention.

The term "specifically hybridize" relates to the ability of a nucleic acid molecule of the present invention to hybridize to at least approximately 6, 12, 20, 30, 50, 100, 150, 200, 300, 350, 400 or 440 successive nucleotides of one of the ß-tubulin genes described above, preferably to one of the sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11 or sequences homologous or complementary thereto, namely in such a way that ten times more molecules hybridize, preferably 100 times more molecules hybridize, and particularly preferably more than 100 times more molecules hybridize than to a cellular nucleic acid (e.g. mRNA or genomic DNA), which codes for a protein other than the ß-tubulin described above.

The term "plasmid" relates to an extrachromosomal genetic element. The original plasmids used for the present invention are either commercially obtainable, freely accessible or can be derived from such plasmids according to known processes.

The term "vector" describes a DNA element which is used for the introduction of exogenous DNA into host cells. A vector contains a nucleotide sequence which codes for one or more polypeptides. Vectors which are able to control the expression of the genes which they contain are designated as "expression vectors".

The term "gene" relates in the scope of the present invention to a nucleic acid comprising an open reading frame which codes for one of the ß-tubulin polypeptides described above. Exon and possible intron sequences are additionally included here.

The term "interact" or "interaction" describes detectable interactions between molecules. The term "binding" is additionally included here.

The term "modulate" relates both to a stimulation and to a suppression or inhibition of a biochemical process. In the context of the present invention, "modulation" means an inhibition or suppression of the interaction between tubulin polypeptides or fragments or subunits thereof, or a stimulation of this interaction which can be shown, for example, in an irreversible binding of tubulin polypeptides to one another.

The term "nucleic acid" relates to polynucleotides such as deoxyribonucleic acids (DNA) or, if appropriate, to ribonucleic acids (RNA). In an equivalent manner, the term also includes analogs of RNA or DNA which are prepared from nucleotide analogs, and, in the case concerned, single-stranded ("sense" or "antisense") and double-stranded polynucleotides.

The term "promoter" relates to DNA sequences which regulate the expression of a specific DNA, which are functionally linked to the promoter. The term also includes "tissue-specific" promoters, i.e. promoters which control the expression of the specific DNA only in certain cells (e.g. cells of a certain tissue). Likewise included are "tissue-nonspecific" promoters and promoters which lead to a constitutive expression or are inducible.

The terms "protein", "polypeptide" and "peptide" are exchangeable in their use in the context of the present application if they relate to a gene product.

A "fusion protein" is a fusion of a first amino acid sequence coding for one of the β-tubulin polypeptides described above having a second amino acid sequence which has no commonality or fundamental homology to the tubulin sequence. The second amino acid sequence can in this case originate from the same organism as the first, or alternatively can originate from another organism (intergenic). In general, a fusion protein can be represented by means of the formula X-tubulin-Y, where "tubulin" represents one of the polypeptides described above, and X and Y represent a polypeptide which is not connected to a tubulin amino acid sequence. X or Y can in each case independently of one another be absent.

The terms "cell" or "host cell" can be used in the same sense in the context of the application present here. It is understood that these terms relate not only to an individual cell, but also to the descendants of such a cell. On account of certain modifications in the course of following generations (e.g. mutations), such descendants are possibly not identical to the stem cell, but are additionally included by the present invention.

The term "intron" describes those sequences of the described, preferably genomic, DNA which are transcribed, but are then removed from the transcript by "splicing", the adjacent sequences (exons) being linked.

Nucleic acids

As already described, one aspect of the invention relates to nucleic acids from nematodes of the family of the Strongylidae, particularly of the subfamily of the Cyathostominae, very particularly to the genera Cyathostomum and Cylicocyclus, especially of the species *Cylicocyclus nassatus* and *Cyathostomum coronatum*, which code for ß-tubulin polypeptides or fragments thereof or nucleic acids homologous thereto which are homologous to 85% to the [lacuna] in SEQ ID NO. 1, 3, 5, 7, 9 and 11, preferably to 95%, and code for a ß-tubulin according to one of the sequences as set forth in SEQ ID NO. 2, 4, 6, 8 or 10 or fragments thereof. SEQ ID NO. 3 represents the degenerated sequence of the nucleic acid from *Cylicocyclus nassatus* coding for β-tubulin, where "r" represents a purine (guanine or adenine), "y" represents a pyrimidine (thymine, or uracil or cytosine) and "w" represents an adenine or a thymine, or a uracil. SEQ ID NO. 3 thus includes a number of sequences which can be present in the organisms of the species *Cylicocyclus nassatus*. The sequences as set forth in SEQ ID NO. 5, 7 and 11 show three defined sequences coding for ß-tubulin, which are exemplary and preferred embodiments of DNA as set forth in SEQ ID NO. 3.

Likewise part of the invention are oligonucleotides which optionally code for ß-tubulin polypentides which comprise a length of at least 2, 5, 10, 25, 50, 100, 150, 200, 250, 300, 350 or 400 amino acids. Such oligonucleotides can be used as primers or antisense molecules (i.e. as noncoding nucleic acids) and comprise at least approximately 6, 12, 24, 30, 60, 100, 120, 150 or 210 base pairs, while coding nucleic acids comprise approximately 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200 or 1300 base pairs.

The invention also describes those oligonucleotides which specifically hybridize under stringent conditions to nucleic acids which are represented by one of the sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11. Correspondingly stringent conditions are, for example, 6 x sodium chloride/sodium citrate (SSC) at approximately 45°C, followed by a washing step with 2 x SSC at 50°C, and are familiar to the person skilled in the art (see, for example, Current protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6). Thus the salt concentrations in the washing step can be chosen such that the stringency is lower (2 x SSC, 50°C) or is higher (0.2 x SSC, 50°C). Furthermore, the temperature in the washing step can be varied, from conditions for a low stringency (e.g. about 22°C), up to conditions of higher stringency (e.g. about 65°C). Both salt concentration and temperature can be varied and tailored to one another.

Particularly preferred oligonucleotides which can be used as primers or for hybridization for the identification and characterization of an existing, e.g. genomic, DNA, are described in SEQ ID NO. 12-51. However, other oligonucleotides can also be derived from the sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11, which can then be used as primers or for hybridization. Particularly preferred for the identification of the species or genus to which an existing DNA can be assigned are those oligonucleotides which hybridize in the region of the intron of an existing (genomic) DNA. The introns of a genomic DNA coding for β-tubulin, which can be isolated from the abovementioned nematodes, are described by way of example of *Cylicocyclus nassatus* in SEQ ID NO. 11. The introns are thus those sequences which

are located between the coding exons. In a preferred embodiment, the described primers or hybridization probes contain a labeled group which makes possible the detection of the oligonucleotides, i.e. for example, radioisotopes, fluorescent groups, enzymes or enzyme cofactors.

Such oligonucleotides can be employed in diagnostic test kits in order to determine the origin (i.e. the organism) of an existing DNA. The oligonucleotides are suitable due to their specific hybridization with the DNAs mentioned under SEQ ID NO. 1, 3, 5, 7, 9 and 11 [lacuna] their fragments, sequences homologous thereto and sequences complementary thereto for recognizing defined sequences. Thus they also make possible the recognition and identification of those sequences coding for ß-tubulin which, on account of one or more point mutations in codon 200, lead to the expression of benzimidazole-resistant ß-tubulin. These oligonucleotides derived from the sequences SEQ ID NO. 1, 3, 5, 7, 9 and 11, preferably the embodiments described under SEQ ID NO. 12 to 51, are thus suitable for the identification of the frequently occurring nematode species of the subfamily of the cyathostominae, and for the recognition of existing resistances to benzimidazoles.

The oligonucleotides according to the present invention can be prepared using standard methods which are familiar to the person skilled in the art, e.g. by de novo DNA synthesis.

The nucleic acids mentioned here can be present in partially purified or biologically pure form in complete cells or in cell lysates, i.e. if other cell components or chemical precursors and by-products have been separated off in the case of a chemical synthesis of the DNA.

Nucleic acids coding for ß-tubulin, as described above, can be obtained starting from mRNA which is present in a number of eukaryotic cells. It is also possible to obtain the DNA according to the invention starting from genomic DNA from the nematode cells concerned (see also the following examples). A gene coding for ß-tubulin can

be obtained, for example, from a cDNA library or a genomic DNA library. cDNA can be obtained by isolating the total mRNA of a cell, e.g. of a nematode cell. Starting from the mRNA, double-stranded cDNA can then be prepared and inserted into a suitable plasmid or a suitable vector. The DNA according to the invention can also be obtained by amplification with the aid of the known polymerase chain reaction (PCR) or alternatively by de novo DNA synthesis (see also J. Sambrook et al. (1989) Molecular Cloning, 2nd Edition, Chap. 14).

Vectors and plasmids

The present invention also comprises expression vectors which contain one of the nucleic acid sequences according to the invention, which are functionally linked to a transcription-regulatory sequence. "Functionally linked" means that the nucleic acid sequence is linked to the regulatory sequence in a manner such that the expression of the protein encoded by the nucleic acid sequence can be controlled. "Transcription-regulatory sequences" include, for example, promoters, enhancers and other control elements. The expression vectors contain, for example, a gene coding for a ß-tubulin according to the invention or fragments thereof. These vectors can be used for incorporation into cells where the corresponding polypeptides or alternatively fusion proteins are then formed. Suitable promoters for the expression of the protein according to the invention in *E.coli* include natural hybrid or bacteriophage promoters. Preferably, they are promoters from the group of the phage λ promoters, omp promoters or synthetic promoters (see also WO 98/15625, DE 3 430 683, EP 0 173 149). Suitable vectors are commercially obtainable, e.g. the expression vectors of the pET series (e.g. pET3a, pET23a, pET28a and His tag or pET32a with His tag) or pGEX with glutathion synthetase fusion. The expression vectors can then be transformed, for example, into DE3-lysogenic E.coli strains, e.g. BL21(DE3), HM S 174(DE3) or AD494(DE3).

Expression of the ß-tubulin polypeptides

The present invention also comprises cells which contain the nucleic acid sequences according to the invention (e.g. inserted into a vector or into the genome). These host cells can be prokaryotic or eukaryotic.

Suitable prokaryotic expression systems are, for example, bacterial systems such as *Streptomyces* spp., *Bacillus subtilis, Salmonella typhimurium, Serratia marcescens* and preferably *E.coli*.

A preferred eukaryotic expression system is the baculovirus system, particularly preferably that which allows post-translational modifications.

Other eukaryotic expression systems (e.g. yeast, insect cells) can likewise be used.

Polypeptides

The present invention likewise comprises ß-tubulin polypeptides which are encoded by the DNAs according to the invention, preferably the DNA sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 and 11, by fragments thereof or by homologous DNA sequences as described above. Preferred embodiments of these ß-tubulin polypeptides are described in the sequences as set forth in SEQ ID NO. 2, 4, 6, 8 and 10. In a preferred embodiment, the described polypeptides are purified polypeptides which are free of contaminating proteins of those cells in which the polypeptides according to the invention have been produced.

The polypeptides described are proteins of full length or fragments, motifs or domains thereof which comprise lengths of at least 5, 10, 25, 50, 75, 100, 125, 150, 200, 250, 300, 350 or 400 amino acids. Polypeptide fragments can be obtained and

are selected by the testing of polypeptides which are encoded by nucleic acid fragments derived from the sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 and 11.

Polypeptide fragments can also be chemically synthesized in a known manner.

The invention also comprises polypeptides which are encoded by the degenerate sequence as set forth in SEQ ID NO. 3. Owing to the various possible bases in defined positions of the DNA sequence, various polypeptides result having an amino acid encoded by the codon resulting in each case. The polypeptides encoded by DNA as set forth in SEQ ID NO. 3 are described in SEQ ID NO. 4, the variable amino acids being marked by "Xaa".

Preferred embodiments of the polypeptides are described in SEQ ID NO. 2, 4, 6, 8 and 10.

The present invention comprises all processes for the preparation of the polypeptides according to the invention.

It is known to the person skilled in the art that the polypeptides of the present invention can be obtained in various ways, e.g. by chemical methods such as the solid-phase method. For the obtainment of larger amounts of protein, the use of recombinant methods is recommended.

The underlying steps for the preparation of the recombinant ß-tubulin are:

1. Obtainment of a natural, synthetic or semisynthetic DNA which codes for the ß-tubulin according to the invention.

2. Incorporation of this DNA into an expression vector which is suitable for expressing the ß-tubulin according to the invention, either on its own or as a fusion protein.

3. Transformation of a suitable, preferably prokaryotic, host cell using this expression vector.

4. Growth of this transformed host cell in a manner which is suitable for expressing the ß-tubulin according to the invention.

5. Harvesting of the cells and purification of the ß-tubulin by means of suitable, known methods.

For example, the expression vectors can be transformed in λDE3-lysogenic *E. coli* strains, e.g. BL21(DE3), HM S174(DE3) or AD494(DE3). After the growth of the cells under the standard conditions familiar to the person skilled in the art, expression is induced using IPTG. After induction of the cells, incubation is carried out at temperatures of 18 to 37°C for 3 to 24 hours. The cells are disrupted, the expressed protein is purified by means of chromatographic methods, in the case of protein expressed with His tag by means of FPLC on an Ni-NTA column, and also by ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis or alternatively immunoaffinity purification, which are specific for the polypeptides according to the invention.

Homologs or fragments of the polypeptides according to the invention can be generated by mutagenesis, such as, for example, by directed (point) mutagenesis, or by deletions.

The polypeptides according to the invention can also be chemically modified, e.g. with glycosyl groups, lipids, phosphates, acetyl groups or similar groups. Covalent derivatives can be obtained by the linkage of the modifying group with functional groups of the amino acid side chains or the N terminus or C terminus of the polypeptide.

In the expression of the polypeptides according to the present invention, it may be advantageous to change certain codons in order to make possible optimum expression. This is true if the use of certain codons ("codon usage") in the heterologous expression system is different than in one of the organisms according to the invention. Furthermore, the deletion of the 5'- or 3'-untranslated region is possible, e.g. if several destabilizing sequence motifs (e.g. ATTTA) are present in the 3' region of the cDNA.

Fusion proteins

The polypeptides according to the invention can also be present as part of a fusion protein. Such fusion proteins are embraced in full by the present invention. Fusion proteins can be useful under conditions where it is desirable to obtain an immunogenic fragment of the ß-tubulin (see, for example, EP 0 259 149; Schlienger et al. (1992) J. Virol. 66, 2). Under certain circumstances, fusion proteins facilitate the expression of a polypeptide. For example, the polypeptides according to the invention can be prepared as glutathione S-transferase (GST) fusion proteins. Such GST fusion proteins facilitate easy purification of the polypeptide (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al. (John Wiley & Sons, N.Y. 1991). Fusion proteins can contain, for example, a "leader" sequence which is used for the purification, e.g. allows a poly-His sequence at the N terminus (but also at the C terminus) of the protein whose purification [lacuna] by means of chromatography on an Ni^{2+} NTA column (see, for example, Hachuli et al. (1987) J. Chromatography 411, 177).

Techniques for preparing such fusion proteins are familiar to the person skilled in the art.

Antibodies

Another aspect of the present invention relates to antibodies which react specifically with the ß-tubulin polypeptides according to the invention.

For example, anti-protein or anti-peptide antisera or monoclonal antibodies can be prepared according to standard protocols by the use of immunogens which have been derived from β-tubulin polypeptides according to the invention (see, for example, Antibodies: A. Laboratory Manual ed. by Harlow and Lane (Cold Spring Harbor Press, 1988)).

Mammals such as mice, hamsters or rabbits can be immunized using an immunogenic form or an immunogenic fraction of the polypeptide according to the invention, e.g. using a polypeptide which is able to produce an antibody response (see also "fusion proteins" above). The appropriate techniques are familiar to the person skilled in the art. Thus an immunogenic fraction of the ß-tubulin can be administered in the presence of an adjuvant. The course of the immunization can be observed by checking the antibody titer in the plasma or serum, e.g. by customary ELISA assays or other immunoassays.

In a preferred embodiment, the antibodies according to the invention are immunospecific for an antigenic determinant of a ß-tubulin polypeptide according to the invention, e.g. of a polypeptide as set forth in SEQ ID NO. 2, 4, 6, 8 or 10 or those polypeptides which are encoded by DNAs as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11 or sequences identical to 85% therewith, preferably sequences identical to 95%.

After the immunization of a mammal, polyclonal anti-ß-tubulin antibodies can be isolated from the serum. For the production of monoclonal antibodies, antibody-producing cells (lymphocytes) can be obtained from an immunized animal and fused

according to known methods with immortal cells such as myeloma cells in order to obtain hybridoma cells (see, for example, Köhler and Milstein (1975) Nature 256, 495-497; Kozbar et al. (1983) Immunology Today 4, 72; Cole et al. (1985) Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. pp. 77-96).

The "antibodies" mentioned here should also include fragments of antibodies which react specifically with ß-tubulin according to the invention. Antibodies can be fragmented using conventional techniques and the fragments tested.

A preferred embodiment relates to antibodies such as described above, which carry a detectable label (e.g. radioisotopes, fluorescent groups, enzymes or enzyme cofactors).

Antibodies which bind specifically to ß-tubulin polypeptides according to the invention can also be used for the immunohistochemical staining of tissue samples in order to detect the expression of a specific β-tubulin. The anti-ß-tubulin antibodies can likewise be used for diagnostic purposes, e.g. for immunoprecipitation or for immunoblotting.

Diagnostic test procedures

The present invention likewise makes available nucleic acid molecules which can be used for diagnostic purposes.

These include nucleic acid molecules such as described above, the fragments of the DNA sequences described under SEQ ID NO. 1, 3, 5, 7, 9 or 11 or DNA sequences complementary thereto. For example, oligonucleotides as set forth in SEQ ID NO. 12 to 51 are made available, which are able to hybridize to sense or antisense sequences coding for ß-tubulin, and to intron sequence sections which are described by way of example in SEQ ID NO. 11.

In this procedure, the nucleic acid of a cell is rendered accessible to hybridization, the DNA probe is brought into contact with the oligonucleotides, and the hybridization of the sample is detected using the oligonucleotide.

In this manner, a method is made available which makes it possible to differentiate between various species of small strongylids and/or other nematode species by means of the specific hybridization of the oligonucleotides according to the invention to a DNA probe, preferably with the aid of those oligonucleotides which hybridize to the intron regions of the DNA coding for ß-tubulin. In a particularly preferred embodiment, the oligonucleotides according to the invention allow the identification of resistances in small (Cyathostaminae) e.g. in horses, especially resistances of the species *Cylicocyclus nassatus*, *Cyathostomum coronatum* and *Cyathostomum catinatum*. In this procedure, the fact can be used that resistant forms of the ß-tubulin carry at least one point mutation in the DNA according to the invention coding therefor, which can be detected by means of PCR, such as described, for example, in an analogous manner in Elard et al. (1998) PCR diagnosis of benzimidazole-susceptibility or –resistance in natural populations of the small ruminant parasite, *Teladorsagia circumcincta*. Veterinary Parasitology 80, 231-237.

The method described is particularly helpful for the assessment of possible treatment strategies in humans and animals affected with nematodes, such as, for example, horses, sheep, pigs, goats, camels, buffalo, donkeys, hares, roe deer, fur-bearing animals, birds (e.g. chickens, turkeys, ducks), fresh- and salt-water fish (e.g. trout, carp). It makes possible the identification and differentiation of the parasitic nematodes and the recognition of resistant populations thereof, and avoid a treatment with inactive nematicides.

The methods described here can be made available, for example, in the form of prefabricated diagnostic test kits which contain at least one of the abovementioned nucleic acid molecules or an antibody such as described above, which is prepared in ready-to-use form.

Procedure for the discovery of nematicidal substances

The invention relates to a procedure in which, with the aid of tubulin or fragments thereof, novel, specific anthelmintic substances can be identified.

In a preferred embodiment, β-tubulin polypeptides according to the present invention are used for this. However, the procedure can also be carried out using tubulin from species other than those mentioned here. Procedures which use β-tubulin polypeptides other than those according to the invention are included in full by the present invention. Particularly preferably, β-tubulin polypeptides as set forth in SEQ ID NO. 2, 4, 6, 8 or 10 are used for the procedure mentioned. In the present invention, recombinant β-tubulin polypeptides from frequently occurring parasitic nematodes are thus additionally made available. These can be used in various test systems for the identification of new inhibitors of tubulin interaction, or the interaction of tubulin subunits.

Cell-free test systems

Many test systems which have the testing of compounds and natural extracts as their aim are aimed at high throughput numbers in order to maximize the number of substances investigated in a given period of time. Test systems which are based on cell-free studies use purified or semi-purified protein. They are suitable for a "first" test, which primarily aims at detecting a possible influence of a substance on the target protein.

Effects such as cell toxicity are as a rule ignored in these *in vitro* systems. The test systems in this case check both inhibitory and suppressive effects of the substances, and stimulatory effects. The effectiveness of a substance can be checked by means of concentration-dependent test series. Control batches without test substances can be used for the assessment of the effects.

One possibility for the identification of substances which modulate the interaction of tubulin or of its subunits is the "scintillation proximity assay" (SPA), see EP 015 473. This test system uses the interaction of a receptor (e.g. tubulin) with a radiolabeled ligand (e.g. a small organic molecule or a second, radiolabeled protein molecule). The receptor is in this case bound to small spheres ("microspheres") or beads which are provided with scintillating molecules. In the course of the decay of the radioactivity, the scintillating substance in the microsphere is excited by the subatomic particles of the radioactive label and a detectable photon is emitted. The test conditions are optimized such that only those particles emanating from the ligand lead to a signal which emanate from a ligand bound to the receptor or the tubulin.

In one possible embodiment, tubulin is bound to the beads, either together or without interacting or binding test substances. α- or β-Tubulin subunits could be used here. A radiolabeled ligand could be, for example, a labeled benzimidazole or a further, labeled β-tubulin molecule. In the case of binding of the ligand to the immobilized tubulin, this ligand must inhibit or abolish an existing interaction between immobilized and free tubulin in order to bind itself in the region of the contact surface. Binding to the immobilized tubulin which has taken place can then be detected by means of a flash of light. Correspondingly, an existing complex between an immobilized and a free, labeled tubulin is destroyed by the binding of a test substance, which leads to a decrease in the detected intensity of the flash of light. The test system then corresponds to a complementary inhibition system.

Test system based on cells

The β-tubulin available by means of the present invention, but also tubulin from other species, makes possible the development of test systems, which are based on cells, for the identification of substances which inhibit the tubulin interaction.

An example of such a test system is the "two hybrid system". A specific example of this is the "interaction trap". This is a genetic selection of interacting proteins in yeast (see, for example, Gyuris et al. (1993) Cdi 1, a human G1 and S phase protein phosphatase that associates with Cdk 2. Cell 75, 791-803). The test system is designed to detect and to describe the interaction of two proteins in that an interaction which has taken place leads to a detectable signal.

Such a test system can also be adapted to the testing of large numbers of test substances in a given period of time.

The system is based on the construction of two vectors, the "bait" vector and the "prey" vector. A gene coding for tubulin, preferably a gene coding for a β-tubulin according to the invention, is cloned in the bait vector and then expressed as a fusion protein with the LexA protein, a DNA-binding protein. A second gene, coding for tubulin, preferably for a β-tubulin according to the invention, is cloned in the prey vector, where it is expressed as a fusion protein with the B42 prey protein. Both vectors are present in a *Saccharomyces cerevisiae* host, which contains copies of a LexA-binding DNA on the 5'-side of a lacZ or HIS3 reporter gene. If an interaction takes place between the two tubulin (fusion) proteins, activation of the transcription of the reporter gene occurs. If the presence of a test substance leads to the inhibition or disturbance of the tubulin interaction, the two tubulin (fusion) proteins can no longer interact and the product of the reporter gene is no longer prepared.

With the aid of tubulin, particularly of the β-tubulin according to the invention or fragments thereof, and the processes described above, it is possible to identify new and specific antiparasitic compounds.

Compounds which are found with the aid of the processes and polypeptides described are valuable for the treatment of humans and animals which are infected with pathogenic endoparasites of the human or of agricultural animals, pets, zoo animals, and laboratory and experimental animals.

The compounds are active against all stages of development of normal, sensitive strains and also resistant strains. By treatment with agents which contain one or more of these compounds, both economic losses in the case of agricultural animals and diseases in humans and animals can be avoided or treated. The following parasites are in this case of particular interest as targets of the active compounds found:

Enoplida, e.g. *Trichuris* spp., *Capillaria* spp., *Trichomosoides* spp., *Trichinella* spp.

Rhabditia, e.g. *Micronema* spp., *Strongyloides* spp.

Strongylida, e.g. *Strongylus* spp., *Triodontophorus* spp., *Oesophagodontus* spp., *Trichonema* spp., *Gyalocephalus* spp., *Cylindropharynx* spp., *Poteriostomum* spp., *Cyclococercus* spp., *Cylicostephanus* spp., *Oesophagostomum* spp., *Chabertia* spp., *Stephanurus* spp., *Ancylostoma* spp., *Uncinaria* spp., *Bunostomum* spp., *Globocephalus* spp., *Syngamus* spp., *Cyathostomum* spp., *Cylicocyclus* spp., *Neostrongylus* spp., *Cystocaulus* spp., *Pneumostrongylus* spp., *Spicocaulus* spp., *Elaphostrongylus* spp., *Parelaphostrongylus* spp., *Crenosoma* spp., *Paracrenosoma* spp., *Angiostrongylus* spp., *Aelurostrongylus* spp., *Filaroides* spp., *Parafilaroides* spp., *Trichostrongylus* spp., *Haemonchus* spp., *Ostertagia* spp., *Marshallagia* spp., *Cooperia* spp., *Nematodirus* spp., *Hyostrongylus* spp., *Obeliscoides* spp., *Amidostomum* spp., *Ollulanus* spp.

Oxyurida, e.g. *Oxyuris* spp., *Enterobius* spp., *Passalurus* spp., *Syphacia* spp., *Aspiculuris* spp., *Heterakis* spp.

Ascaridia, e.g. *Ascaris* spp., *Toxascaris* spp., *Toxocara* spp., Parascaris spp., *Anisakis* spp., *Ascaridia* spp.

Spirurida, e.g. *Gnathostoma* spp., *Physaloptera* spp., *Thelazia* spp., *Gongylonema* spp., *Habronema* spp., Parabronema spp., *Draschia* spp., *Dracunculus* spp.

Filariida, e.g. *Stephanofilaria* spp., *Parafilaria* spp., *Setaria* spp., *Loa* spp., *Dirofilaria* spp., *Litomosoides* spp., *Brugia* spp., *Wuchereria* spp., *Onchocerca* spp.

Gigantorhynchida, e.g. *Filicollis* spp., *Moniliformis* spp., *Macracanthorhynchus* spp., *Prosthenorchis* spp.

Mastigophora (Flagellata)

Trypanosomatidae, e.g. *Trypanosoma b. brucei, T. b. gambiense, T. b. rhodesiense, T. congolense, T. cruzi, T. evansi, T. equinum, T. lewisi, T. percae, T. simiae, T. vivax, Leishmania brasiliensis, L. donovani, L. tropica*

Trichomonadidae , e.g. *Giardia lambilia, G. canis.*

Sarcomastigophora (Rhizopoda), e.g. *Entamoeba histolytica*

Hartmanellidae, e.g. *Acanthamoeba* sp., *Hartmanella* spp.

Apicomplexa (Sporozoa), e.g. *Eimeria acervulina, E. adenoides, E. alabahmensis, E. anatis, E. anseris, E. arloingi, E. ashata, E. auburnensis, E. bovis, E. brunetti, E. canis, E. chinchillae, E. clupearum, E. columbae, E. contorta, E. crandalis, E. debliecki, E. dispersa, E. ellipsoidales, E. falciformis, E. faurei, E. labbeana, E. leucarti, E. magna, E. maxima, E. media, E. meleagridis. E. meleagrimitis, E. mitis, E. necatrix, E. ninakohlyakimovae, E. ovis, E. parva, E. pavonis, E. perforans, E. phasani, E. piriformis, E. praecox, E. residua, E. scabra, E.* spec., *E. stiedai, E. suis, E. tenella, E. truncata, E. truttae, E. zuernii, Globidium* spec., *Isospora belli, I. canis, I. felis, I. ohioensis, I. rivolta, I.* spec., *I. suis, Neospora caninum, Cystisospora* spec. *Cryptosporidium* spec.

Toxoplasmadidae, e.g. *Toxoplasma gondii*

Sarcocystidae, e.g. *Sarcocystis bovicanis, S. bovihominis, S. neuvona, S. ovicanis, S. ovifelis, S.* spec., *S. suihominis*

Leucozoide, e.g. *Leucozytozoon simondi*

Plasmodiidae, e.g. *Plasmodium berghei, P. falciparum, P. malariae, P. ovale, P. vivax, P.* spec.

Piroplasmea, e.g. *Babesia argentina, B. bovis, B. canis, B.* spec., *Theileria parva, T.* spec.

Adeleina, e.g. *Hepatozoon canis, H.* spec.

The following are furthermore of importance

Myxospora and Microspora, e.g. *Glugea* spec. and *Nosema* spec., and *Pneumocystis carinii*, Ciliophora (Ciliata), e.g. *Balantidium coli, Ichthiophthirius* spec., *Trichondina* spec. or *Epistylis* spec.

The compounds and agents found are likewise effective against protozoa of insects, such as those of the strain Microsporidia, particularly those of the order Nosema, very particularly those of the species Nosema apis, which are parasites of honeybees.

Examples

Example 1

Obtainment of β-tubulin cDNA and genomic DNA

mRNA was obtained with the aid of the Quick Prep® Micro mRNA kit (Pharmacia Biotech, Freiburg, Germany) from *C. nassatus* worms and from *C. coronatum* and *C.catinatum* worms using the Dynal® mRNA Direct Kit (Dynal, Hamburg, Germany). The worms were isolated from the large intestine of horses and differentiated microscopically according to the characterized structure of head and tail (see R. S. Lichtenfeld (1975), Helminths of domestic equids. Proceedings of the Helminthological Society, Washington, 42 (Special issue), 1-92).

The synthesis of the cDNA was carried out with the aid of the "reverse transcription system" (Promega, Madison, USA) in the case of the mRNA from *C. nassatus* and using the superscript RTII reverse transcriptase (Gibco BRL Life Technologies) in the case of the mRNA from *C. coronatum* and *C. catinatum*. In the cases mentioned, oligonucleotides having a length of 15 base pairs were used. Incubation was carried out for one hour at 42°C.

Genomic DNA was obtained from 4 to 40 adult worms using the QIA Amp-tissue kit (Qiagen, Hilden, Germany). In this procedure, the worms were digested for 2 hours at 55°C with proteinase K and the genomic DNA extracted using "spin columns".

Example 2

Amplification of β-tubulin sequences

The amplification of β-tubulin sequences of full length or of fragments can be carried out, for example, using AmpliTaq Gold™ polymerase (Perkin Elmer, Foster City, California, USA).

An amplification of the β-tubulin sequences as set forth in SEQ ID NO. 1, 3, 5, 7, 9 or 11 or fragments thereof can be carried out with the aid of the primers as set forth in SEQ ID NO. 12 - 51.

For the amplification of the cDNA from *C. coronatum*, for example, the sequences as set forth in SEQ ID NO. 43 and 44 are suitable, for the amplification of the cDNA from *C. catinatum*, the sequences as set forth in SEQ ID NO. 40 and 42 are particularly suitable.

The amplification of the *C. nassatus* cDNA and of the genomic DNA of all species according to the present invention was carried out in a total volume of 50 µl, containing 5 µl of 10 x buffer, 2.5 µl of $MgCl_2$ (25 mM), 2 µl of dNTP mix (2mM for each NTP), 1 µl of each specific primer (SEQ ID NO. 12 - 47) (50 p mol/µl), 0.5 µl (2.5 U) of polymerase and 1 - 5 µl of DNA template. When using degenerate primers (SEQ ID NO. 48 - 51), 1 µl of each primer of a conc. of 500 pmol/µl was employed. The annealing was carried out in the case of degenerate primers at 46°C, in the case of specific primers the temperature was varied according to the calculated melting temperature. The PCR cycles were chosen as follows:

95°C for 10 min, then 35 - 40 cycles with 1 min denaturation at 94°C, 1 min annealing, 1 min at 72°C and a final step at 72°C for 10 min. In the amplification of cDNA from *C. coronatum* and *C. catinatum*, a "touchdown" PCR temperature program was carried out, which has the following profile:

firstly 15 cycles at 94°C for 30 sec, then 1 min at 60°C and 1 min at 72°C, followed by 15 cycles with 30 sec at 95°C, 55°C for 1 min and 72°C for 1 min and finally 10 cycles at 95°C for 30 sec, then 45°C for 1 min and 72°C for 1 min. For the amplification of larger fragments (>1000 base pairs), the elongation phase at 72°C was lengthened to 2.30 min.

Example 3

PCR products from the amplification of cDNA or genomic DNA from *C. nassatus, C. coronatum* and *C. catinatum* were cloned with the aid of the "Original TA Cloning Kit" /Invitrogen, Leek, Netherlands), namely in the "Original TA Cloning®" vector.

Patent claims

1. A DNA coding for β-tubulin from Cyathostominae or fragments thereof.

2. A DNA as claimed in claim 1, comprising

   a) a polynucleotide having at least 85% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 2;

   b) a polynucleotide having at least 85% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 4;

   c) a polynucleotide having at least 85% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 6;

   d) a polynucleotide having at least 85% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 8;

   e) a polynucleotide having at least 85% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 10.

3. A DNA as claimed in claim 1, comprising

   a) a polynucleotide having at least 95% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 2;

   b) a polynucleotide having at least 95% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 4;

   c) a polynucleotide having at least 95% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 6;

d) a polynucleotide having at least 95% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 8;

e) a polynucleotide having at least 95% identity to a polynucleotide coding for an amino acid sequence as set forth in SEQ ID NO. 10.

4. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 1.

5. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 3.

6. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 5.

7. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 7.

8. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 9.

9. A DNA as claimed in one of claims 1 to 3, comprising a sequence as set forth in SEQ ID NO. 11.

10. A DNA as claimed in one of claims 1 to 3 and 5 to 9, characterized in that it originates from Cylicocyclus.

11. A DNA as claimed in one of claims 1 to 4, characterized in that it originates from Cyathostomum.

12. A DNA as claimed in one of claims 1 to 3 and 5 to 10, characterized in that it originates from *Cylicocyclus nassatus*.

13. A DNA as claimed in one of claims 1 to 4 and 11, characterized in that it originates from *Cyathostomum coronatum*.

14. A DNA as claimed in one of claims 1 to 13, characterized in that it contains at least one base replacement in codon 200, which leads to the expression of a polypeptide having anthelmintic resistance.

15. A DNA, characterized in that it is complementary to DNA as claimed in one of claims 1 to 14 or fragments thereof.

16. An RNA, characterized in that it is complementary to DNA as claimed in one of claims 1 to 15.

17. An expression construct, characterized in that it comprises DNA as claimed in one of claims 1 to 14 and a sequence linked functionally therewith, which makes possible the expression of the DNA.

18. A vector, characterized in that it comprises DNA as claimed in one of claims 1 to 14.

19. A host cell, comprising DNA as claimed in one of claims 1 to 14, an expression construct as claimed in claim 17, or a vector as claimed in claim 18.

20. A polypeptide encoded by a DNA as claimed in one of claims 1 to 14 or fragments thereof.

21. A polypeptide as claimed in claim 20, consisting of or comprising an amino acid sequence as set forth in SEQ ID NO. 2.

22. A polypeptide as claimed in claim 20, consisting of or comprising an amino acid sequence as set forth in SEQ ID NO. 4.

23. A polypeptide as claimed in claim 20, consisting of or comprising an amino acid sequence as set forth in SEQ ID NO. 6.

24. A polypeptide as claimed in claim 20, consisting of or comprising an amino acid sequence as set forth in SEQ ID NO. 8.

25. A polypeptide as claimed in claim 20, consisting of or comprising an amino acid sequence as set forth in SEQ ID NO. 10.

26. A polypeptide encoded by a DNA as claimed in claim 14.

27. A process for the preparation of a polypeptide as claimed in one of claims 20 to 26, comprising the expression of the polypeptide or fragments thereof in a prokaryotic or eukaryotic expression system.

28. The use of DNA oligonucleotides which hybridize specifically to DNA as claimed in one of claims 1 to 15, preferably to noncoding DNA sections, for the detection of DNA which originates from Cyathostominae.

29. The use of DNA which hybridizes specifically to DNA as claimed in one of claims 1 to 15 for the detection of DNA which originates from Cyathostominae and codes for a polypeptide as claimed in claim 26.

30. A procedure for the detection of Cyathostominae, characterized in that DNA as set forth in claim 28 is hybridized to DNA as claimed in one of claims 1 to 15 and this is amplified by means of PCR.

31. A procedure for the detection of Cyathostominae having anthelmintic resistance, characterized in that DNA as set forth in claim 29 is hybridized to DNA as claimed in one of claims 1 to 15 and this is amplified by means of PCR.

32. A DNA oligonucleotide comprising one of the sequences as set forth in SEQ ID NO. 12 to SEQ ID NO. 51 or a sequence derived from one of the DNA sequences as claimed in claims 1 to 15.

33. A diagnostic test kit comprising at least one of the oligonucleotides as claimed in claim 32 and/or antibodies as claimed in claim 35 or 36.

34. A diagnostic test kit as claimed in claim 33, characterized in that the DNA oligonucleotides are provided with a detectable label.

35. An antibody, characterized in that it reacts specifically with an epitope of a polypeptide as claimed in one of claims 20 to 26.

36. An antibody as claimed in claim 35, characterized in that it is monoclonal.

37. The use of antibodies as claimed in claim 35 or 36 as nematicides.

38. The use of polypeptides as claimed in one of claims 20 to 26 for the production of vaccines.

39. A procedure for the identification of substances which modulate the interaction of tubulin.

40. The procedure as claimed in claim 39, characterized in that

a) the test substance is brought into contact with tubulin under those conditions which allow interaction of the tubulin molecules with one another and binding of the test substance to tubulin,

b) the binding of the test substance which has taken place is detected by determining the ability of the tubulin protein molecules to interact with one another and

c) the ability of the tubulin protein molecules to interact with one another in the presence of a test substance is compared with their ability to interact with one another in the absence of a test substance.

41. The procedure as claimed in claim 39 or 40, characterized in that the tubulin used is a polypeptide as claimed in one of claims 20 to 26.

42. The procedure as claimed in one of claims 39 to 41, characterized in that, for the detection of a modulation of the tubulin interaction in the presence of a test substance, a test system based on cells is used.

43. The procedure as claimed in one of claims 39 to 41, characterized in that, for the detection of a modulation of the tubulin interaction in the presence of a test substance, a cell-free test system is used.

44. A substance which is identified in a procedure as claimed in one of claims 39 to 43.

45. The use of a substance as claimed in claim 44 for the production of an agent for the prophylactic or therapeutic treatment of nematode attack.

DNA coding for β-tubulin and its use

Abstract

The invention relates to DNA which codes for β-tubulin from nematodes of the family of the strongylidae, the polypeptide encoded by this DNA, the use of the DNA for the diagnosis of anthelmintic resistance in the nematodes concerned and for the identification of the species of parasitic nematodes, the use of the β-tubulin as a constituent of a vaccine, and a process for the identification of new anthelmintic or antibiotic compounds.

COMBINED DECLARATION AND POWER OF ATTORNEY	ATTORNEY DOCKET NO

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name. I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought
on the invention entitled

DNA CODING FOR ß-TUBULIN AND USE THEREOF

the specification of which is attached hereto,

or was filed on **June 30, 2000**

as a PCT Application Serial No. **PCT/EP00/06104**

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s), the priority(ies) of which is/are to be claimed:

199 31 883.2	**Germany**	**July 9, 1999**
(Number)	(Country)	(Month/Day/Year Filed)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose the material information as defined in Title 37, Code of Federal Regulations, §1.56 which occured between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

10030566 .D10703

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and this application and to transact all business in the Patent and Trademark Office conne

JOSEPH C. GIL, Patent Office Registration Number 26,602 ARON PREIS, Patent Office Registration Number 29,4
LYNDANNE M. WHALEN, Patent Office Registration Number 29,457 THOMAS W. ROY,
Patent Office Registration Number 29,582 RICHARD E. L. HENDERSON, Patent Office Registration Number 31,619
GODFRIED R. AKORLI, Patent Office Registration Number 28,779 N. DENISE BROWN, Patent Office
Registration Number 36,097 NOLAND J. CHEUNG, Patent Office Registration Number 39,138
DIDERICO VAN EYL, Patent Office Registration Number 38,641 CAROLYN M. SLOANE, Patent Office
Registration Number 44,339 JAMES R. FRANKS, Patent Office Registration Number 42,552
JACKIE ANN ZURCHER, Patent Office Registration Number 42,251
RAYMOND J. HARMUTH, Patent Office Registration Number 33,896
all of Bayer Corporation, Pittsburgh, Pennsylvania 15205-9741

Send Correspondence To: Direct Telephone Calls To:
Patent Department
Bayer Corporation (412) 777-2349
100 Bayer Road
Pittsburgh, Pennsylvania 15205-9741

Customer No. 27941

	FULL NAME OF SOLE OR FIRST INVENTOR	INVENTOR'S SIGNATURE	DATE
1-00	Georg von Samson-Himmelstjerna		23.11.2001
	RESIDENCE	CITIZENSHIP	
	D 42657 Solingen, Germany DEX	German	
	POST OFFICE ADDRESS		
	Neuenkamper Str. 21, D 42657 Solingen, Germany		
	FULL NAME OF SECOND INVENTOR	INVENTOR'S SIGNATURE	DATE
2-00	Achim Harder		07.11.19
	RESIDENCE	CITIZENSHIP	
	D 51109 Köln, Germany DEX	German	
	POST OFFICE ADDRESS		
	c/o BAYER AKTIENGESELLSCHAFT, D 51368 Leverkusen, Germany		
	FULL NAME OF THIRD INVENTOR	INVENTOR'S SIGNATURE	DATE
3-00	Thomas Schnieder		23. Nov. 2001
	RESIDENCE	CITIZENSHIP	
	D 30539 Hannover, Germany DEX	German	
	POST OFFICE ADDRESS		
	In der Bebie 38, D 30539 Hannover, Germany		
	FULL NAME OF FOURTH INVENTOR	INVENTOR'S SIGNATURE	DATE
4-00	Michaela Pape		22.11.01
	RESIDENCE	CITIZENSHIP	
	D 30173 Hannover, Germany DEX	German	
	POST OFFICE ADDRESS		
	Elkart Allee 21, D 30173 Hannover, Germany		
	FULL NAME OF FIFTH INVENTOR	INVENTOR'S SIGNATURE	DATE
	RESIDENCE	CITIZENSHIP	
	POST OFFICE ADDRESS		
	FULL NAME OF SIXTH INVENTOR	INVENTOR'S SIGNATURE	DATE
	RESIDENCE	CITIZENSHIP	
	POST OFFICE ADDRESS		
	FULL NAME OF SEVENTH INVENTOR	INVENTOR'S SIGNATURE	DATE
	RESIDENCE	CITIZENSHIP	
	POST OFFICE ADDRESS		

10030566.010702

PTO/SB/122 (10-01)
Approved for use through 10/31/2002. OMB 0651-0035
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

CHANGE OF CORRESPONDENCE ADDRESS
Application

Address to:
Assistant Commissioner for Patents
Washington, D.C. 20231

Application Number	**10/030,566**
Filing Date	**01/07/2002**
First Named Inventor	**Von Samson-Himmelstjerna**
Art Unit	
Examiner Name	
Attorney Docket Number	**MO-6878/LeA 33,759**

Please change the Correspondence Address for the above-identified application to:

[X] Customer Number **27941** → *Place Customer Number Bar Code Label here*

Type Customer Number here

OR

[ ] Firm *or* Individual Name	
Address	
Address	
City	**State** **ZIP**
Country	
Telephone	**Fax**

This form cannot be used to change the data associated with a Customer Number. To change the data associated with an existing Customer Number use "Request for Customer Number Data Change" (PTO/SB/124).

CERTIFICATE OF MAILING UNDER 37 CFR 1.8
The undersigned hereby certifies that this paper and any papers referred to as attached were being deposited with the United States Postal Service, with sufficient postage, as first class mail in an envelope addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231 on this 30 day of July, 2002

Signature of Person Certifying

I am the :

[ ] Applicant/Inventor.

[ ] Assignee of record of the entire interest.
Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96).

[X] Attorney or Agent of record.

[ ] Registered practitioner named in the application transmittal letter in an application without an executed oath or declaration. See 37 CFR 1.33(a)(1). Registration Number _____

Typed or Printed Name **Susan M. Pellegrino**

Signature *Susan M. Pellegrino*

Date **July 30, 2002**

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

[X] *Total of **1** forms are submitted.

Burden Hour Statement: This form is estimated to take 3 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

10/030566 #4.

PTO/SB/122 (10-01)
Approved for use through 10/31/2002. OMB 0651-0035
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

CHANGE OF CORRESPONDENCE ADDRESS
Application

Address to:
Assistant Commissioner for Patents
Washington, D.C. 20231

Application Number	**10/030,566**
Filing Date	**01/07/2002**
First Named Inventor	**Von Samson-Himmelstjerna**
Art Unit	
Examiner Name	
Attorney Docket Number	**MO-6878/LeA 33,759**

Please change the Correspondence Address for the above-identified application to:

[X] Customer Number **27941** ——→

Type Customer Number here

Place Customer Number Bar Code Label here

OR

[ ] **Firm** *or* **Individual Name**					
Address					
Address					
City		**State**		**ZIP**	
Country					
Telephone		**Fax**			

This form cannot be used to change the data associated with a Customer Number. To change the data associated with an existing Customer Number use "Request for Customer Number Data Change" (PTO/SB/124).

I am the :

[ ] Applicant/Inventor.

[ ] Assignee of record of the entire interest.
Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96).

[X] Attorney or Agent of record.

[ ] Registered practitioner named in the application transmittal letter in an application without an executed oath or declaration. See 37 CFR 1.33(a)(1). Registration Number _____

Typed or Printed Name **Susan M. Pellegrino**

Signature *Susan M. Pellegrino*

Date **July 30, 2002**

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

[X] *Total of **1** forms are submitted.

Burden Hour Statement: This form is estimated to take 3 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

PATENT APPLICATION
Mo-6878
LeA 33,759

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION OF)
)
GEORG VON SAMSON-HIMMELSTJERNA ET AL)
)
SERIAL NUMBER: 10/030,566)
)
FILED: January 7, 2002)
)
TITLE: DNA CODING FOR β-TUBULIN AND USE THEREOF)

ASSOCIATE POWER OF ATTORNEY

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

I hereby appoint:

Jeffrey M. Greenman Reg. No. 26,552
Alice A. Brewer, Reg. No. 32,888
Jerrie L. Chiu, Reg. No. 41,670
William F. Gray, Reg. No. 31,018
Susan M. Pellegrino, Reg. No. 48,972
Barbara A. Shimei, Reg. No. 29,862

whose post office address is: Bayer Corporation, 400 Morgan Lane, West Haven, CT 06516-4175, as Attorneys of Record in this application, to make alterations and amendments therein and to transact all business in the Patent and Trademark Office connected therewith.

Respectfully submitted,

By _____
Godfried R. Akorli
Attorney for Applicants
Reg. No. 28,779

COPY #4.

PATENT APPLICATION
Mo-6878
LeA 33,759

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION OF)
)
GEORG VON SAMSON-HIMMELSTJERNA)
ET AL)
)
SERIAL NUMBER: 10/030,566)
)
FILED: January 7, 2002)
)
TITLE: DNA CODING FOR β-TUBULIN)
AND USE THEREOF)

ASSOCIATE POWER OF ATTORNEY

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

I hereby appoint:

Jeffrey M. Greenman Reg. No. 26,552
Alice A. Brewer, Reg. No. 32,888
Jerrie L. Chiu, Reg. No. 41,670
William F. Gray, Reg. No. 31,018
Susan M. Pellegrino, Reg. No. 48,972
Barbara A. Shimei, Reg. No. 29,862

whose post office address is: Bayer Corporation, 400 Morgan Lane, West Haven, CT 06516-4175, as Attorneys of Record in this application, to make alterations and amendments therein and to transact all business in the Patent and Trademark Office connected therewith.

Respectfully submitted,

By _____
Godfried R. Akorli
Attorney for Applicants
Reg. No. 28,779

SEQUENCE LISTING

<110> Bayer AG

<120> DNA coding for β-tubulin and its use

<130> Le A 33 759

<140>
<141>

<160> 51

<170> PatentIn Ver. 2.1

<210> 1
<211> 1380
<212> DNA
<213> Cyathostomum coronatum

<220>
<221> CDS
<222> (1)..(1344)

<400> 1

atg cgt gag atc gtg cat gta caa gct gga caa tgt gga aac caa att 48
Met Arg Glu Ile Val His Val Gln Ala Gly Gln Cys Gly Asn Gln Ile
1 5 10 15

ggt tcc aag ttt tgg gaa gtg atc tct gac gag cat ggc att aag ccc 96
Gly Ser Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Lys Pro

20 25 30

gat ggc aca tac cac gga gaa tct gat cta caa tta gaa cga atc aat 144
Asp Gly Thr Tyr His Gly Glu Ser Asp Leu Gln Leu Glu Arg Ile Asn
35 40 45

gtg tac tat aat gaa gca cat gga ggc aaa tat gtc cca cgt gca gtt 192
Val Tyr Tyr Asn Glu Ala His Gly Gly Lys Tyr Val Pro Arg Ala Val
50 55 60

ctt gtt gat ctc gag ccc gga act atg gat tcc gtc cgt tcc ggg cca 240
Leu Val Asp Leu Glu Pro Gly Thr Met Asp Ser Val Arg Ser Gly Pro
65 70 75 80

tac ggg caa ttg ttc cgc cct gat aac tac gtg ttt gga cag tct ggc 288
Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr Val Phe Gly Gln Ser Gly
85 90 95

gca gga aat aac tgg gca aaa ggt cac tac act gaa ggc gct gaa ctt 336
Ala Gly Asn Asn Trp Ala Lys Gly His Tyr Thr Glu Gly Ala Glu Leu
100 105 110

gtc gac aat gta cta gat gta gtg cga aaa gaa gca gaa gga tgt gac 384
Val Asp Asn Val Leu Asp Val Val Arg Lys Glu Ala Glu Gly Cys Asp
115 120 125

tgt ctg cag ggc ttc cag cta act cac tca ctt gga gga ggt acc ggt 432
Cys Leu Gln Gly Phe Gln Leu Thr His Ser Leu Gly Gly Gly Thr Gly
130 135 140

tcg ggt atg ggc act ctc ctc atc tcc aaa att cgg gag gag tat cct 480

```
Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu Tyr Pro
145                 150                 155                 160

gat aga atc atg tcc tcg ttc tcc gtt gtc ccc tca cca aag gtc tcc   528
Asp Arg Ile Met Ser Ser Phe Ser Val Val Pro Ser Pro Lys Val Ser
                165                 170                 175

gac act gtt gtg gag cct tac aat gct acc cta tcc gtt cat cag ttg   576
Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser Val His Gln Leu
            180                 185                 190

gtt gaa aat aca gac gag act tat tgt att gac aat gaa gcc ctg tat   624
Val Glu Asn Thr Asp Glu Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr
        195                 200                 205

gat att tgc ttc cgc act ttg aaa ctc acg aac cca act tat gga gat   672
Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr Asn Pro Thr Tyr Gly Asp
    210                 215                 220

ctg aat cat ctt gtg tct gta aca atg tct ggt gtc acc aca tgt ctt   720
Leu Asn His Leu Val Ser Val Thr Met Ser Gly Val Thr Thr Cys Leu
225                 230                 235                 240

cgc ttc cct ggc caa ttg aat gcc gat cta cgc aaa cta gct gtt aac   768
Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu Arg Lys Leu Ala Val Asn
                245                 250                 255

atg gtt cca ttc cct cgt ctt cac ttc ttc atg cct ggt ttt gct cct   816
Met Val Pro Phe Pro Arg Leu His Phe Phe Met Pro Gly Phe Ala Pro
            260                 265                 270
```

```
ctt tct gct aaa ggt gct cag gct tac cgt gct ctt acc gta gcc gag   864
Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg Ala Leu Thr Val Ala Glu
        275                 280                 285

ctt aca cag cag atg ttt gat gct aag aat atg atg gct gct tgc gac   912
Leu Thr Gln Gln Met Phe Asp Ala Lys Asn Met Met Ala Ala Cys Asp
    290                 295                 300

cct cga cat gga cgt tat ctc acc gtc gca gcc atg ttc cga gga aga   960
Pro Arg His Gly Arg Tyr Leu Thr Val Ala Ala Met Phe Arg Gly Arg
305                 310                 315                 320

atg agc atg agg gaa gta gac gac cag atg atg tca gtg cag aac aag
1008
Met Ser Met Arg Glu Val Asp Asp Gln Met Met Ser Val Gln Asn Lys
                325                 330                 335

aac tcc tca tac ttc gta gag tgg atc ccg aac aac gtg aag acc gct
1056
Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala
            340                 345                 350

gta tgc gac atc ccg cca cga gga ctg aag atg gcc gct acc ttc gtt
1104
Val Cys Asp Ile Pro Pro Arg Gly Leu Lys Met Ala Ala Thr Phe Val
        355                 360                 365

gga aac tca act gcc atc caa gag ctg ttc aag cgc att tca gaa caa
1152
Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser Glu Gln
    370                 375                 380

ttt aca gcc atg ttc cgc cgc aaa gcg ttc ttg cat tgg tac act ggt
1200
```

Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr Thr Gly
385 390 395 400

gaa ggt atg gac gag atg gag ttc act gaa gca gag tcc aac atg aat
1248

Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn Met Asn
405 410 415

gat ctc atc tcc gag tac caa cag tac cag gaa gcc acc gct gac gac
1296

Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln Glu Ala Thr Ala Asp Asp
420 425 430

atg ggc gat ctt gat gcg gaa ggc gct gaa gag gct tat cct gag gaa
1344

Met Gly Asp Leu Asp Ala Glu Gly Ala Glu Glu Ala Tyr Pro Glu Glu
435 440 445

taaaccagca gatcgtgttg cgttgttcgt ttctct
1380

<210> 2

<211> 448

<212> PRT

<213> Cyathostomum coronatum

<400> 2

Met Arg Glu Ile Val His Val Gln Ala Gly Gln Cys Gly Asn Gln Ile
1 5 10 15

Gly Ser Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Lys Pro
20 25 30

Asp Gly Thr Tyr His Gly Glu Ser Asp Leu Gln Leu Glu Arg Ile Asn
35 40 45

Val Tyr Tyr Asn Glu Ala His Gly Gly Lys Tyr Val Pro Arg Ala Val
50 55 60

Leu Val Asp Leu Glu Pro Gly Thr Met Asp Ser Val Arg Ser Gly Pro
65 70 75 80

Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr Val Phe Gly Gln Ser Gly
85 90 95

Ala Gly Asn Asn Trp Ala Lys Gly His Tyr Thr Glu Gly Ala Glu Leu
100 105 110

Val Asp Asn Val Leu Asp Val Val Arg Lys Glu Ala Glu Gly Cys Asp
115 120 125

Cys Leu Gln Gly Phe Gln Leu Thr His Ser Leu Gly Gly Gly Thr Gly
130 135 140

Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu Tyr Pro
145 150 155 160

Asp Arg Ile Met Ser Ser Phe Ser Val Val Pro Ser Pro Lys Val Ser
165 170 175

Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser Val His Gln Leu
180 185 190

Val Glu Asn Thr Asp Glu Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr
195 200 205

Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr Asn Pro Thr Tyr Gly Asp
210 215 220

Leu Asn His Leu Val Ser Val Thr Met Ser Gly Val Thr Thr Cys Leu
225 230 235 240

Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu Arg Lys Leu Ala Val Asn
245 250 255

Met Val Pro Phe Pro Arg Leu His Phe Phe Met Pro Gly Phe Ala Pro
260 265 270

Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg Ala Leu Thr Val Ala Glu
275 280 285

Leu Thr Gln Gln Met Phe Asp Ala Lys Asn Met Met Ala Ala Cys Asp
290 295 300

Pro Arg His Gly Arg Tyr Leu Thr Val Ala Ala Met Phe Arg Gly Arg
305 310 315 320

Met Ser Met Arg Glu Val Asp Asp Gln Met Met Ser Val Gln Asn Lys
325 330 335

Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala
340 345 350

Val Cys Asp Ile Pro Pro Arg Gly Leu Lys Met Ala Ala Thr Phe Val

355 360 365

Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser Glu Gln

370 375 380

Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr Thr Gly

385 390 395 400

Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn Met Asn

405 410 415

Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln Glu Ala Thr Ala Asp Asp

420 425 430

Met Gly Asp Leu Asp Ala Glu Gly Ala Glu Glu Ala Tyr Pro Glu Glu

435 440 445

<210> 3

<211> 1429

<212> DNA

<213> Cylicocyclus nassatus

<220>

<221> CDS

<222> (1)..(1362)

<400> 3

aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga 48

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly

```
1               5                  10                  15

car tgt gga aac caa att ggy tcc aag tty tgg gaa gtg atc tct gac   96
Gln Cys Gly Asn Gln Ile Xaa Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

gag cac ggc att aag ccy gay ggc aca tac cay gga gaa tct gay yta   144
Glu His Gly Ile Lys Xaa Asp Gly Thr Tyr His Gly Glu Ser Asp Xaa
        35                  40                  45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aar   192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat   240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

tcr gtc cgy tcy ggg cca tac ggg caa ttg ttc cgc cct gat aac tac   288
Xaa Val Xaa Xaa Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac   336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

act gaa ggy gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa   384
Thr Glu Xaa Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca   432
```

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

ctt gga gga ggt acc gga tcg rgt atg ggc acw ctc ctc atc tyc aaa 480
Leu Gly Gly Gly Thr Gly Ser Xaa Met Gly Xaa Leu Leu Ile Xaa Lys
145 150 155 160

att cgg gag gag tat cct gat aga atc atr tcc tcg ttc tyc gtt gtt 528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Xaa Ser Ser Phe Xaa Val Val
165 170 175

ccc tca cca aag gtc tyc gay acy gtt gtg gag ccg tac aat gct acc 576
Pro Ser Pro Lys Val Xaa Asp Xaa Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gar act twc tgt att 624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Xaa Cys Ile
195 200 205

gac aat gaa gct ctt tat gat att tgc ttc cgc acy ytg aaa ctc acs 672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Xaa Xaa Lys Leu Xaa
210 215 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct 720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

ggy gtc act aca tgy ctt cgc ttc cct ggc caa ttg rry gcc gat ctw 768
Xaa Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Xaa Ala Asp Xaa
245 250 255

cgt aaa cta gct gtt aac atg gyt cca ttc cct cgt ctt cac tty tty 816
Arg Lys Leu Ala Val Asn Met Xaa Pro Phe Pro Arg Leu His Phe Phe
260 265 270

atg cct ggc ttt gct ccc ctc tct gcy aaa ggc gcy cag gct tac cgt 864
Met Pro Gly Phe Ala Pro Leu Ser Xaa Lys Gly Xaa Gln Ala Tyr Arg
275 280 285

gct ctt act gta gcc gag ctw acy caa yag atg ttc gat gcc aaa aat 912
Ala Leu Thr Val Ala Glu Xaa Xaa Gln Xaa Met Phe Asp Ala Lys Asn
290 295 300

atg atg gcc gct tgc gac cct cga cat gga crt tat ctc acc gty gca 960
Met Met Ala Ala Cys Asp Pro Arg His Gly Xaa Tyr Leu Thr Xaa Ala
305 310 315 320

gcc atg ttc cga gga cga atg agc ayg agg gar gta gac gac cag atg
1008
Ala Met Phe Arg Gly Arg Met Ser Xaa Arg Glu Val Asp Asp Gln Met
325 330 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg
1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

aac aac gtc aar acc gcy gta tgc gac att ccg ccr aga gga ctg aaa
1104
Asn Asn Val Lys Thr Xaa Val Cys Asp Ile Pro Xaa Arg Gly Leu Lys
355 360 365

atg gcc gct acc ttc gtt gga aac yca act gcc aty caa gag ctg tty
1152

Met Ala Ala Thr Phe Val Gly Asn Xaa Thr Ala Xaa Gln Glu Leu Phe

370 375 380

aag cgc att tca gaa caa tty aca gct atg ttc cgc cgc aaa gcg tty
1200

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe

385 390 395 400

ttg cat ygg tay act ggw gaa ggt atg gay gag atg gag ttc act gaa
1248

Leu His Xaa Tyr Thr Xaa Glu Gly Met Asp Glu Met Glu Phe Thr Glu

405 410 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac car caa tac cag
1296

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln

420 425 430

gaa gct acm gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa
1344

Glu Ala Xaa Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu

435 440 445

gag gct tac cct gar gaa tagamcagca gaytgtgttg cgttgttcgt
1392

Glu Ala Tyr Pro Glu Glu

450

ttctctrtgt caatgcgaaa tacacattga ttgcgtt
1429

<210> 4

<211> 454

<212> PRT

<213> Cylicocyclus nassatus

<400> 4

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

Gln Cys Gly Asn Gln Ile Xaa Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

Glu His Gly Ile Lys Xaa Asp Gly Thr Tyr His Gly Glu Ser Asp Xaa
35 40 45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

Xaa Val Xaa Xaa Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

Thr Glu Xaa Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

Leu Gly Gly Gly Thr Gly Ser Xaa Met Gly Xaa Leu Leu Ile Xaa Lys
145 150 155 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Xaa Ser Ser Phe Xaa Val Val
165 170 175

Pro Ser Pro Lys Val Xaa Asp Xaa Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Xaa Cys Ile
195 200 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Xaa Xaa Lys Leu Xaa
210 215 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

Xaa Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Xaa Ala Asp Xaa
245 250 255

Arg Lys Leu Ala Val Asn Met Xaa Pro Phe Pro Arg Leu His Phe Phe
260 265 270

Met Pro Gly Phe Ala Pro Leu Ser Xaa Lys Gly Xaa Gln Ala Tyr Arg
275 280 285

Ala Leu Thr Val Ala Glu Xaa Xaa Gln Xaa Met Phe Asp Ala Lys Asn
290 295 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Xaa Tyr Leu Thr Xaa Ala

305 310 315 320

Ala Met Phe Arg Gly Arg Met Ser Xaa Arg Glu Val Asp Asp Gln Met
325 330 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

Asn Asn Val Lys Thr Xaa Val Cys Asp Ile Pro Xaa Arg Gly Leu Lys
355 360 365

Met Ala Ala Thr Phe Val Gly Asn Xaa Thr Ala Xaa Gln Glu Leu Phe
370 375 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385 390 395 400

Leu His Xaa Tyr Thr Xaa Glu Gly Met Asp Glu Met Glu Phe Thr Glu
405 410 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
420 425 430

Glu Ala Xaa Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
435 440 445

Glu Ala Tyr Pro Glu Glu
450

<210> 5

<211> 1428

<212> DNA

<213> Cylicocyclus nassatus

<220>

<221> CDS

<222> (1)..(1362)

<400> 5

aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga 48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

cag tgt gga aac caa att ggc tcc aag ttt tgg gaa gtg atc tct gac 96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

gag cac ggc att aag cct gat ggc aca tac cac gga gaa tct gat tta 144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
35 40 45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aaa 192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat 240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

tcg gtc cgt tcc ggg cca tac ggg caa ttg ttc cgc cct gat aac tac 288

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac 336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

act gaa ggc gct gaa ctt gtt gac aat gta cta gat gta gtg cga aaa 384
Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca 432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

ctt gga gga ggt acc gga tcg ggt atg ggc act ctc ctc atc tcc aaa 480
Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145 150 155 160

att cgg gag gag tat cct gat aga atc atg tcc tcg ttc tcc gtt gtt 528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
165 170 175

ccc tca cca aag gtc tcc gac acc gtt gtg gag ccg tac aat gct acc 576
Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att 624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

```
gac aat gaa gct ctt tat gat att tgc ttc cgc act ttg aaa ctc acg   672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct   720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

ggt gtc act aca tgt ctt cgc ttc cct ggc caa ttg aat gcc gat cta   768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
                245                 250                 255

cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttt   816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

atg cct ggc ttt gct ccc ctc tct gct aaa ggc gct cag gct tac cgt   864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat   912
Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca   960
Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg
1008
Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335
```

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg
1056

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro

340 345 350

aac aac gtc aag acc gct gta tgc gac att ccg ccg aga gga ctg aaa
1104

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys

355 360 365

atg gcc gct acc ttc gtt gga aac tca act gcc atc caa gag ctg ttc
1152

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe

370 375 380

aag cgc att tca gaa caa ttc aca gct atg ttc cgc cgc aaa gcg ttc
1200

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe

385 390 395 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa
1248

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu

405 410 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac cag caa tac cag
1296

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln

420 425 430

gaa gct aca gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa
1344

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu

435 440 445

gag gct tac cct gaa gaa tagacagcag attgtgttgc gttgttcgtt
1392

Glu Ala Tyr Pro Glu Glu

450

tctctgtgtc aatgcgaaat acacattgat tgcgtt
1428

<210> 6

<211> 454

<212> PRT

<213> Cylicocyclus nassatus

<400> 6

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
35 40 45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr

85 90 95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145 150 155 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
165 170 175

Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
210 215 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
245 250 255

Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
260 265 270

Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
275 280 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
290 295 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305 310 315 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
325 330 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
355 360 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
370 375 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385 390 395 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
405 410 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
420 425 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
435 440 445

Glu Ala Tyr Pro Glu Glu
450

<210> 7
<211> 1429
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> CDS
<222> (1)..(1362)

<400> 7
aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga 48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

cag tgt gga aac caa att ggt tcc aag ttt tgg gaa gtg atc tct gac 96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

gag cac ggc att aag ccc gat ggc aca tac cac gga gaa tct gac tta 144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu

35 40 45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aaa 192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat 240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

tcg gtc cgt tcc ggg cca tac ggg caa ttg ttc cgc cct gat aac tac 288
Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac 336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

act gaa ggc gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa 384
Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca 432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

ctt gga gga ggt acc gga tcg agt atg ggc act ctc ctc atc ttc aaa 480
Leu Gly Gly Gly Thr Gly Ser Ser Met Gly Thr Leu Leu Ile Phe Lys
145 150 155 160

att cgg gag gag tat cct gat aga atc ata tcc tcg ttc ttc gtt gtt 528

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Ile Ser Ser Phe Phe Val Val
165 170 175

ccc tca cca aag gtc tcc gac acc gtt gtg gag ccg tac aat gct acc 576
Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att 624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

gac aat gaa gct ctt tat gat att tgc ttc cgc act ttg aaa ctc acg 672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
210 215 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct 720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

ggt gtc act aca tgt ctt cgc ttc cct ggc caa ttg agt gcc gat cta 768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Ser Ala Asp Leu
245 250 255

cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttt 816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
260 265 270

atg cct ggc ttt gct ccc ctc tct gct aaa ggc gct cag gct tac cgt 864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
275 280 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat 912

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn

290 295 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca 960

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala

305 310 315 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg
1008

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met

325 330 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg
1056

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro

340 345 350

aac aac gtc aag acc gct gta tgc gac att ccg ccg aga gga ctg aaa
1104

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys

355 360 365

atg gcc gct acc ttc gtt gga aac tca act gcc att caa gag ctg ttc
1152

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe

370 375 380

aag cgc att tca gaa caa ttt aca gct atg ttc cgc cgc aaa gcg ttc
1200

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe

385 390 395 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa
1248

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
405 410 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac caa caa tac cag
1296

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
420 425 430

gaa gct acc gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa
1344

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
435 440 445

gag gct tac cct gag gaa tagaacagca gattgtgttg cgttgttcgt
1392

Glu Ala Tyr Pro Glu Glu
450

ttctctgtgt caatgcgaaa tacacattga ttgcgtt
1429

<210> 8

<211> 454

<212> PRT

<213> Cylicocyclus nassatus

<400> 8

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp

20 25 30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
35 40 45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

Leu Gly Gly Gly Thr Gly Ser Ser Met Gly Thr Leu Leu Ile Phe Lys
145 150 155 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Ile Ser Ser Phe Phe Val Val
165 170 175

Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
210 215 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Ser Ala Asp Leu
245 250 255

Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
260 265 270

Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
275 280 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
290 295 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305 310 315 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
325 330 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
355 360 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
370 375 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385 390 395 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
405 410 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
420 425 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
435 440 445

Glu Ala Tyr Pro Glu Glu
450

<210> 9
<211> 1428
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> CDS
<222> (1)..(1362)

<400> 9

aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga 48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

cag tgt gga aac caa att ggt tcc aag ttc tgg gaa gtg atc tct gac 96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

gag cac ggc att aag ccc gac ggc aca tac cat gga gaa tct gat cta 144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
35 40 45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aag 192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat 240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

tca gtc cgt tct ggg cca tac ggg caa ttg ttc cgc cct gat aac tac 288
Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac 336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

act gaa ggc gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa 384

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca 432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

ctt gga gga ggt acc gga tcg ggt atg ggc aca ctc ctc atc tcc aaa 480
Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145 150 155 160

att cgg gag gag tat cct gat aga atc atg tcc tcg ttc tcc gtt gtt 528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
165 170 175

ccc tca cca aag gtc ttc gat act gtt gtg gag ccg tac aat gct acc 576
Pro Ser Pro Lys Val Phe Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att 624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

gac aat gaa gct ctt tat gat att tgc ttc cgc acc ttg aaa ctc acg 672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
210 215 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct 720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

ggt gtc act aca tgc ctt cgc ttc cct ggc caa ttg aat gcc gat cta 768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
245 250 255

cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttc 816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
260 265 270

atg cct ggc ttt gct ccc ctc tct gcc aaa ggc gcc cag gct tac cgt 864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
275 280 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat 912
Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
290 295 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca 960
Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305 310 315 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg 1008
Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
325 330 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg 1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

aac aac gtc aaa acc gct gta tgc gac att ccg ccg aga gga ctg aaa 1104
Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys

355 360 365

atg gcc gct acc ttc gtt gga aac tca act gcc att caa gag ctg ttc
1152

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe

370 375 380

aag cgc att tca gaa caa ttc aca gct atg ttc cgc cgc aaa gcg ttc
1200

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe

385 390 395 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa
1248

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu

405 410 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac cag caa tac cag
1296

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln

420 425 430

gaa gct acc gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa
1344

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu

435 440 445

gag gct tac cct gaa gaa tagacagcag attgtgttgc gttgttcgtt
1392

Glu Ala Tyr Pro Glu Glu

450

tctctgtgtc aatgcgaaat acacattgat tgcgtt
1428

<210> 10

<211> 454

<212> PRT

<213> Cylicocyclus nassatus

<400> 10

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
35 40 45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145 150 155 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
165 170 175

Pro Ser Pro Lys Val Phe Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
195 200 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
210 215 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
245 250 255

Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
260 265 270

Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
275 280 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn

290 295 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305 310 315 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
325 330 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
340 345 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
355 360 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
370 375 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385 390 395 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
405 410 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
420 425 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
435 440 445

Glu Ala Tyr Pro Glu Glu
450

<210> 11

<211> 2655

<212> DNA

<213> Cylicocyclus nassatus

<220>

<221> intron

<222> (1)..(18)

<220>

<221> intron

<222> (1)..(18)

<220>

<221> intron

<222> (76)..(358)

<220>

<221> intron

<222> (469)..(637)

<220>

<221> intron

<222> (865)..(1374)

<220>

<221> intron

<222> (1666)..(1723)

<220>

<221> intron

<222> (1915)..(1966)

<220>

<221> intron

<222> (2064)..(2119)

<220>

<221> intron

<222> (2306)..(2354)

<220>

<221> intron

<222> (2475)..(2523)

<220>

<221> intron

<222> (2592)..(2655)

<400> 11

aagttctcta ctgcaataat gcgtgagatc gtgcatgtac aagctggaca rtgtggaaac 60

caaattggtt ccaaggtrcg gtagtttyrr twktytrytg atcgtaattc sggmgytytr 120

dagtrrytttt ttycgytgyy ratgttgcat yrtgttgcga taaascyraa aawtcawwag 180

rcgaggctgt aaaagsactt ytacttttra atmcrytgta gcagcatgag tcatcrgcat 240

gtttgcagtg sgtttttttat gcgcwgawcc yctcagaaga tgagaatgcg wtccaytgag 300

cwtagartct grctttctct cgttawctaa ratcaamtta carcrytyca ttttkcagtt 360

ytgggaagtg atctctgacg agcacggcat taagccygay ggcacatacc ayggagaatc 420

tgatytacaa ttagaacgaa tcaatgtgta ctataatgaa gcacatggtt agtcgtayat 480

ccgcttcgtt gtytcccmat gcagrccyct tcagttttta taactgycga aatatcgatc 540

gggctctttt gcagcggccw ygattacgca ataccayygc ygcygcagtg gcrgtcgaaa 600

ttaatgtggt caracgtgaa aatgtggtgc tttyaggagg caartatgtc ccgcgtgcag 660

ttcttgttga tctcgagccc ggaactatgg attcggtccg ytccgggcca tacgggcaat 720

tgttccgccc tgataactac gtgtttggac agtctggcgc aggaaataac tgggcaaaag 780

gtcactacac tgaaggygct gaacttgtcg acaatgtact agatgtagtg cgaaaagaag 840

ctgaaggatg tgactgtctg caggtaaatt tccaagtagt agcaggaaat ggtwtgtgra 900

tagcataaca aaagtcatag aaggaatatg gacgctagtc aaaacaaagw tggacgttar 960

tcggtcgtcc gggacarttt ggaagtcatg ggtcasccaa cacgcttttt tamaagtaca
1020

tcatactctt ttcccacgaa aagctatttt gcgtattacg gggtacaggg gaggggtcaa
1080

aatcacagat tgctgaaaty tggttcactg ragttattgr tgaaaatcat attgattttg
1140

cttgctactg ccttcttttr aggctatgct ttacaatctt ggggcctgga taaccgaatt
1200

gtcygaagtt tttcggtcat cacggacggg gaaggggcat artatcgtta kttcttgkta
1260

tttcgcagca tatggcaatc tytccacttc tgacaagttt tcygtagaaa atatwcttca
1320

aggtstcaag aacyttgctg ctagrgctgt aaaccaayct gtatcycttt cagggcttcc
1380

agctaactca ctcacttgga ggaggtaccg gatcgggtat gggcactctc ctcatctcca
1440

aaattcggga ggagtatcct gatagaatca tgtcctcgtt ctccgttgtt ccctcaccaa
1500

aggtctccga caccgttgtg gagccgtaca atgctaccct atccgttcat cagttggttg
1560

aaaatacaga cgaracttwc tgtattgaca atgaagctct ttatgatatt tgcttccgca
1620

cyytgaaact cacsaaccca acttatggag atctgaatca tcttggtrrg yrayatkcsa
1680

ytgctgagct tdgtrgaatt tvctaattwt ktyhamtdty yagtgtctgt aacaatgtct
1740

ggygtcacta catgycttcg cttccctggc caattgrayg ccgatctwcg taaactagct
1800

gttaacatgg ytccattccc tcgtcttcac ttyttyatgc ctggctttgc tcccctctct
1860

gcyaaaggcg cycaggctta ccgtgctctt actgtagccg agctwacyca rcaggtgcgt
1920

ctgcttatcr ttgwtgayrt gtgtttattc kttgtrtatt ttayagatgt tcgatgccaa
1980

aaatatgatg gccgcttgcg accctcgaca tggacrttat ctcaccgtyg cagccatgtt
2040

ccgaggacga atgagcayga gggtaagtgg mtkmttggyc ytytaryaya rctcrgacga
2100

awtgctgtta tgtcmtagga rgtagacgac cagatgatgt cagtgcagaa caagaactcc
2160

tcatacttcg tagagtggat tccgaacaac gtcaaraccg cygtatgcga cattccgccr
2220

agaggactga aaatggccgc taccttcgtt ggaaacycaa ctgccatcca agagctgtty
2280

aagcgcattt cagaacaatt yacaggttdg tttgtgcaya ttatggtgaa agcagattar
2340

ttgcgaygtt gcagctatgt tccgccgcaa agcgttyttg catyggtaya ctggwgaagg
2400

tatggaygag atggagttca ctgaagccga gtccaacatg aatgatctca tctccgaata
2460

ccarcaatac caggttcggc tgtytttcwt rgayactgtr tttaataatt wttyttgtct
2520

aggaagctac cgctgacgat atgggcgatc tcgatgcgga aggcgctgaa gaggcttacc
2580

ctgargaata gamcagcaga ytgtgttgcg ttgttcgttt ctctrtgtca atgcgaaata
2640

cacattgatt gcgtt
2655

<210> 12

<211> 23

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: hybridization probe/primer

<400> 12

ggtttaatta tcccaagttt gag 23

<210> 13

<211> 20

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 13

ggccacgcgt cgactagtac 20

<210> 14

<211> 37

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 14

ggccacgcgt cgactagtac ttttttttttt ttttttt 37

<210> 15

<211> 22

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 15

gaccgctgta tgcgacattc cg 22

<210> 16

<211> 22

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 16

aactcaactg ccatccaaga gc 22

<210> 17

<211> 21

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic

sequence: primer/hybridization probes

<400> 17

gctatgttcc gccgcaaagc g 21

<210> 18

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 18

acgagcacgg cattaagcct gatgg 25

<210> 19

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 19

ccatcaggct taatgccgtg ctcgt 25

<210> 20

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 20

ccgaatccat agttccgggc tcgag 25

<210> 21

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 21

ccgacaccgt tgtggagccg taca 24

<210> 22

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 22

gcgaccctcg acatggacgt tatct 25

<210> 23

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 23

agataacgtc catgtcgagg gtcgc 25

<210> 24

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 24

tgttccgagg acgaatgagc atgag 25

<210> 25

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 25

ctcatgctca ttcgtcctcg gaaca 25

<210> 26

<211> 26

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 26

aggtagacga ccagatgatg tcagtg 26

<210> 27

<211> 26

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 27

cactgacatc atctggtcgt ctacct 26

<210> 28

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 28

ggcggaatgt cgcatacagc ggtc 24

<210> 29

<211> 26

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 29
cggagatgag atcattcatg ttggac 26

<210> 30
<211> 24
<212> DNA
<213> Synthetic sequence

<220>
<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 30
ctctcctcat ctccaaaatt cggg 24

<210> 31
<211> 25
<212> DNA
<213> Synthetic sequence

<220>
<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 31
cagctaactc actcacttgg aggag 25

<210> 32

<211> 25

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 32

aagttctcta ctgcaataat gcgtg 25

<210> 33

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 33

ggttgaaaat acagacgaga cttt 24

<210> 34

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 34

ggttgaaaat acagacgaga ctta 24

<210> 35

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 35

aaagagcttc attgtcaata caga 24

<210> 36

<211> 23

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 36

caattggcca gggaagcgaa gac 23

<210> 37

<211> 23

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 37

gtcttcgctt ccctggccaa ttg 23

<210> 38

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic sequence: primer/hybridization probes

<400> 38

cctggctttg ctcccctctc tgct 24

<210> 39

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 39

agcagagagg ggagcaaagc cagg 24

<210> 40

<211> 23

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 40

aacgcaatca atgtgtattt cgc 23

<210> 41

<211> 18

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 41

cccgacggca cataccat 18

<210> 42

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 42

gaaacgaaca acgcaacaca atct 24

<210> 43

<211> 22

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 43

caagctggac aatgtggaaa cc 22

<210> 44

<211> 22

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 44

yagagaaacg aacaacgcaa ca 22

<210> 45

<211> 23

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 45

ttgatctcga gcccggaact atg 23

<210> 46

<211> 22

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 46

tctcgagccc ggaactatgg at 22

<210> 47

<211> 24

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 47

cccgaatttt ggagatgagg agag 24

<210> 48

<211> 20

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic

sequence: primer/hybridization probes

<400> 48

ggncartgyg gnaaycarat 20

<210> 49

<211> 20

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 49

gaytgyytnc arggnttyca 20

<210> 50

<211> 20

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 50

tgraanccyt gnarrcartc 20

<210> 51

<211> 20

<212> DNA

<213> Synthetic sequence

<220>

<223> Description of the synthetic
sequence: primer/hybridization probes

<400> 51

tcytgrtayt gytgrtaytc 20

SEQUENCE LISTING

<110> Von Samson-Himmelstjerna, Georg
Harder, Achim
Schneider, Thomas
Pape, Michaela

<120> DNA Coding for Beta-Tubulin and Use Thereof

<130> Mo6878/LeA 33 759

<140> US 10/030,566
<141> 2002-01-07

<150> PCT/EP00/06104
<151> 2000-06-30

<150> DE19931883.2
<151> 1999-07-09

<160> 2

<170> PatentIn version 3.1

<210> 1
<211> 1380
<212> DNA
<213> Cyathostomum coronatum

<220>
<221> CDS
<222> (1)..(1344)
<223>

<400> 1

```
atg cgt gag atc gtg cat gta caa gct gga caa tgt gga aac caa att       48
Met Arg Glu Ile Val His Val Gln Ala Gly Gln Cys Gly Asn Gln Ile
1               5                  10                  15

ggt tcc aag ttt tgg gaa gtg atc tct gac gag cat ggc att aag ccc       96
Gly Ser Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Lys Pro
            20                  25                  30

gat ggc aca tac cac gga gaa tct gat cta caa tta gaa cga atc aat      144
Asp Gly Thr Tyr His Gly Glu Ser Asp Leu Gln Leu Glu Arg Ile Asn
        35                  40                  45

gtg tac tat aat gaa gca cat gga ggc aaa tat gtc cca cgt gca gtt      192
Val Tyr Tyr Asn Glu Ala His Gly Gly Lys Tyr Val Pro Arg Ala Val
    50                  55                  60

ctt gtt gat ctc gag ccc gga act atg gat tcc gtc cgt tcc ggg cca      240
Leu Val Asp Leu Glu Pro Gly Thr Met Asp Ser Val Arg Ser Gly Pro
65                  70                  75                  80

tac ggg caa ttg ttc cgc cct gat aac tac gtg ttt gga cag tct ggc      288
Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr Val Phe Gly Gln Ser Gly
                85                  90                  95

gca gga aat aac tgg gca aaa ggt cac tac act gaa ggc gct gaa ctt      336
Ala Gly Asn Asn Trp Ala Lys Gly His Tyr Thr Glu Gly Ala Glu Leu
            100                 105                 110
```

```
gtc gac aat gta cta gat gta gtg cga aaa gaa gca gaa gga tgt gac      384
Val Asp Asn Val Leu Asp Val Val Arg Lys Glu Ala Glu Gly Cys Asp
        115                 120                 125

tgt ctg cag ggc ttc cag cta act cac tca ctt gga gga ggt acc ggt      432
Cys Leu Gln Gly Phe Gln Leu Thr His Ser Leu Gly Gly Gly Thr Gly
    130                 135                 140

tcg ggt atg ggc act ctc ctc atc tcc aaa att cgg gag gag tat cct      480
Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu Tyr Pro
145                 150                 155                 160

gat aga atc atg tcc tcg ttc tcc gtt gtc ccc tca cca aag gtc tcc      528
Asp Arg Ile Met Ser Ser Phe Ser Val Val Pro Ser Pro Lys Val Ser
                165                 170                 175

gac act gtt gtg gag cct tac aat gct acc cta tcc gtt cat cag ttg      576
Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser Val His Gln Leu
            180                 185                 190

gtt gaa aat aca gac gag act tat tgt att gac aat gaa gcc ctg tat      624
Val Glu Asn Thr Asp Glu Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr
        195                 200                 205

gat att tgc ttc cgc act ttg aaa ctc acg aac cca act tat gga gat      672
Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr Asn Pro Thr Tyr Gly Asp
    210                 215                 220

ctg aat cat ctt gtg tct gta aca atg tct ggt gtc acc aca tgt ctt      720
Leu Asn His Leu Val Ser Val Thr Met Ser Gly Val Thr Thr Cys Leu
225                 230                 235                 240

cgc ttc cct ggc caa ttg aat gcc gat cta cgc aaa cta gct gtt aac      768
Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu Arg Lys Leu Ala Val Asn
                245                 250                 255

atg gtt cca ttc cct cgt ctt cac ttc ttc atg cct ggt ttt gct cct      816
Met Val Pro Phe Pro Arg Leu His Phe Phe Met Pro Gly Phe Ala Pro
            260                 265                 270

ctt tct gct aaa ggt gct cag gct tac cgt gct ctt acc gta gcc gag      864
Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg Ala Leu Thr Val Ala Glu
        275                 280                 285

ctt aca cag cag atg ttt gat gct aag aat atg atg gct gct tgc gac      912
Leu Thr Gln Gln Met Phe Asp Ala Lys Asn Met Met Ala Ala Cys Asp
    290                 295                 300

cct cga cat gga cgt tat ctc acc gtc gca gcc atg ttc cga gga aga      960
Pro Arg His Gly Arg Tyr Leu Thr Val Ala Ala Met Phe Arg Gly Arg
305                 310                 315                 320

atg agc atg agg gaa gta gac gac cag atg atg tca gtg cag aac aag     1008
Met Ser Met Arg Glu Val Asp Asp Gln Met Met Ser Val Gln Asn Lys
                325                 330                 335

aac tcc tca tac ttc gta gag tgg atc ccg aac aac gtg aag acc gct     1056
Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala
            340                 345                 350

gta tgc gac atc ccg cca cga gga ctg aag atg gcc gct acc ttc gtt     1104
Val Cys Asp Ile Pro Pro Arg Gly Leu Lys Met Ala Ala Thr Phe Val
        355                 360                 365
```

```
gga aac tca act gcc atc caa gag ctg ttc aag cgc att tca gaa caa      1152
Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser Glu Gln
    370                 375                 380

ttt aca gcc atg ttc cgc cgc aaa gcg ttc ttg cat tgg tac act ggt      1200
Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr Thr Gly
385                 390                 395                 400

gaa ggt atg gac gag atg gag ttc act gaa gca gag tcc aac atg aat      1248
Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn Met Asn
                405                 410                 415

gat ctc atc tcc gag tac caa cag tac cag gaa gcc acc gct gac gac      1296
Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln Glu Ala Thr Ala Asp Asp
            420                 425                 430

atg ggc gat ctt gat gcg gaa ggc gct gaa gag gct tat cct gag gaa      1344
Met Gly Asp Leu Asp Ala Glu Gly Ala Glu Glu Ala Tyr Pro Glu Glu
        435                 440                 445

taaaccagca gatcgtgttg cgttgttcgt ttctct                             1380
```

<210> 2
<211> 448
<212> PRT
<213> Cyathostomum coronatum

<400> 2

```
Met Arg Glu Ile Val His Val Gln Ala Gly Gln Cys Gly Asn Gln Ile
1               5                  10                  15

Gly Ser Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Lys Pro
            20                  25                  30

Asp Gly Thr Tyr His Gly Glu Ser Asp Leu Gln Leu Glu Arg Ile Asn
        35                  40                  45

Val Tyr Tyr Asn Glu Ala His Gly Gly Lys Tyr Val Pro Arg Ala Val
    50                  55                  60

Leu Val Asp Leu Glu Pro Gly Thr Met Asp Ser Val Arg Ser Gly Pro
65                  70                  75                  80

Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr Val Phe Gly Gln Ser Gly
                85                  90                  95

Ala Gly Asn Asn Trp Ala Lys Gly His Tyr Thr Glu Gly Ala Glu Leu
            100                 105                 110

Val Asp Asn Val Leu Asp Val Val Arg Lys Glu Ala Glu Gly Cys Asp
        115                 120                 125

Cys Leu Gln Gly Phe Gln Leu Thr His Ser Leu Gly Gly Gly Thr Gly
    130                 135                 140
```

```
Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu Tyr Pro
145                 150                 155                 160

Asp Arg Ile Met Ser Ser Phe Ser Val Val Pro Ser Pro Lys Val Ser
                165                 170                 175

Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser Val His Gln Leu
            180                 185                 190

Val Glu Asn Thr Asp Glu Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr
        195                 200                 205

Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr Asn Pro Thr Tyr Gly Asp
    210                 215                 220

Leu Asn His Leu Val Ser Val Thr Met Ser Gly Val Thr Thr Cys Leu
225                 230                 235                 240

Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu Arg Lys Leu Ala Val Asn
                245                 250                 255

Met Val Pro Phe Pro Arg Leu His Phe Phe Met Pro Gly Phe Ala Pro
            260                 265                 270

Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg Ala Leu Thr Val Ala Glu
        275                 280                 285

Leu Thr Gln Gln Met Phe Asp Ala Lys Asn Met Met Ala Ala Cys Asp
    290                 295                 300

Pro Arg His Gly Arg Tyr Leu Thr Val Ala Ala Met Phe Arg Gly Arg
305                 310                 315                 320

Met Ser Met Arg Glu Val Asp Asp Gln Met Met Ser Val Gln Asn Lys
                325                 330                 335

Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala
            340                 345                 350

Val Cys Asp Ile Pro Pro Arg Gly Leu Lys Met Ala Ala Thr Phe Val
        355                 360                 365

Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser Glu Gln
    370                 375                 380

Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr Thr Gly
385                 390                 395                 400
```

```
Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn Met Asn
                405                 410                 415

Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln Glu Ala Thr Ala Asp Asp
            420                 425                 430

Met Gly Asp Leu Asp Ala Glu Gly Ala Glu Glu Ala Tyr Pro Glu Glu
        435                 440                 445
```

<210> 3
<211> 1429
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> CDS
<222> (1)..(1362)
<223> Xaa is a variable amino acid.

<400> 3

```
aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga        48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15

car tgt gga aac caa att ggy tcc aag tty tgg gaa gtg atc tct gac        96
Gln Cys Gly Asn Gln Ile Xaa Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

gag cac ggc att aag ccy gay ggc aca tac cay gga gaa tct gay yta       144
Glu His Gly Ile Lys Xaa Asp Gly Thr Tyr His Gly Glu Ser Asp Xaa
        35                  40                  45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aar       192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat       240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

tcr gtc cgy tcy ggg cca tac ggg caa ttg ttc cgc cct gat aac tac       288
Xaa Val Arg Xaa Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac       336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

act gaa ggy gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa       384
Thr Glu Xaa Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca       432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140
```

```
ctt gga gga ggt acc gga tcg rgt atg ggc acw ctc ctc atc tyc aaa      480
Leu Gly Gly Gly Thr Gly Ser Xaa Met Gly Xaa Leu Leu Ile Xaa Lys
145                 150                 155                 160

att cgg gag gag tat cct gat aga atc atr tcc tcg ttc tyc gtt gtt      528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Xaa Ser Ser Phe Xaa Val Val
                165                 170                 175

ccc tca cca aag gtc tyc gay acy gtt gtg gag ccg tac aat gct acc      576
Pro Ser Pro Lys Val Xaa Asp Xaa Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gar act twc tgt att      624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Xaa Cys Ile
        195                 200                 205

gac aat gaa gct ctt tat gat att tgc ttc cgc acy ytg aaa ctc acs      672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Xaa Xaa Lys Leu Xaa
    210                 215                 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct      720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

ggy gtc act aca tgy ctt cgc ttc cct ggc caa ttg rry gcc gat ctw      768
Xaa Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Xaa Ala Asp Xaa
                245                 250                 255

cgt aaa cta gct gtt aac atg gyt cca ttc cct cgt ctt cac tty tty      816
Arg Lys Leu Ala Val Asn Met Xaa Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

atg cct ggc ttt gct ccc ctc tct gcy aaa ggc gcy cag gct tac cgt      864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

gct ctt act gta gcc gag ctw acy caa yag atg ttc gat gcc aaa aat      912
Ala Leu Thr Val Ala Glu Xaa Xaa Gln Xaa Met Phe Asp Ala Lys Asn
    290                 295                 300

atg atg gcc gct tgc gac cct cga cat gga crt tat ctc acc gty gca      960
Met Met Ala Ala Cys Asp Pro Arg His Gly Xaa Tyr Leu Thr Xaa Ala
305                 310                 315                 320

gcc atg ttc cga gga cga atg agc ayg agg gar gta gac gac cag atg     1008
Ala Met Phe Arg Gly Arg Met Ser Xaa Arg Glu Val Asp Asp Gln Met
                325                 330                 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg     1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

aac aac gtc aar acc gcy gta tgc gac att ccg ccr aga gga ctg aaa     1104
Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Xaa Arg Gly Leu Lys
        355                 360                 365

atg gcc gct acc ttc gtt gga aac yca act gcc aty caa gag ctg tty     1152
Met Ala Ala Thr Phe Val Gly Asn Xaa Thr Ala Xaa Gln Glu Leu Phe
    370                 375                 380

aag cgc att tca gaa caa tty aca gct atg ttc cgc cgc aaa gcg tty     1200
Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400
```

```
ttg cat ygg tay act ggw gaa ggt atg gay gag atg gag ttc act gaa      1248
Leu His Xaa Tyr Thr Xaa Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac car caa tac cag      1296
Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

gaa gct acm gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa      1344
Glu Ala Xaa Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

gag gct tac cct gar gaa tagamcagca gaytgtgttg cgttgttcgt             1392
Glu Ala Tyr Pro Glu Glu
    450

ttctctrtgt caatgcgaaa tacacattga ttgcgtt                             1429
```

<210> 4
<211> 454
<212> PRT
<213> Cylicocyclus nassatus

<220>
<221> misc_feature
<222> (23)..(23)
<223> The 'Xaa' at location 23 stands for Gly.

<220>
<221> misc_feature
<222> (38)..(38)
<223> The 'Xaa' at location 38 stands for Pro.

<220>
<221> misc_feature
<222> (48)..(48)
<223> The 'Xaa' at location 48 stands for Leu.

<220>
<221> misc_feature
<222> (81)..(81)
<223> The 'Xaa' at location 81 stands for Ser.

<220>
<221> misc_feature
<222> (84)..(84)
<223> The 'Xaa' at location 84 stands for Ser.

<220>
<221> misc_feature
<222> (115)..(115)
<223> The 'Xaa' at location 115 stands for Gly.

<220>
<221> misc_feature
<222> (152)..(152)
<223> The 'Xaa' at location 152 stands for Gly, or Ser.

<220>
<221> misc_feature
<222> (155)..(155)
<223> The 'Xaa' at location 155 stands for Thr.

<220>
<221> misc_feature
<222> (159)..(159)
<223> The 'Xaa' at location 159 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (170)..(170)
<223> The 'Xaa' at location 170 stands for Met, or Ile.

<220>
<221> misc_feature
<222> (174)..(174)
<223> The 'Xaa' at location 174 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (182)..(182)
<223> The 'Xaa' at location 182 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (184)..(184)
<223> The 'Xaa' at location 184 stands for Thr.

<220>
<221> misc_feature
<222> (206)..(206)
<223> The 'Xaa' at location 206 stands for Tyr, or Phe.

<220>
<221> misc_feature
<222> (220)..(220)
<223> The 'Xaa' at location 220 stands for Thr.

<220>
<221> misc_feature
<222> (221)..(221)
<223> The 'Xaa' at location 221 stands for Leu.

<220>
<221> misc_feature
<222> (224)..(224)
<223> The 'Xaa' at location 224 stands for Thr.

<220>
<221> misc_feature
<222> (241)..(241)
<223> The 'Xaa' at location 241 stands for Gly.

<220>
<221> misc_feature
<222> (253)..(253)
<223> The 'Xaa' at location 253 stands for Gly, Asp, Ser, or Asn.

<220>
<221> misc_feature
<222> (256)..(256)
<223> The 'Xaa' at location 256 stands for Leu.

<220>
<221> misc_feature
<222> (264)..(264)
<223> The 'Xaa' at location 264 stands for Ala, or Val.

<220>
<221> misc_feature
<222> (295)..(295)
<223> The 'Xaa' at location 295 stands for Leu.

<220>
<221> misc_feature
<222> (296)..(296)
<223> The 'Xaa' at location 296 stands for Thr.

<220>
<221> misc_feature
<222> (298)..(298)
<223> The 'Xaa' at location 298 stands for Gln.

<220>
<221> misc_feature
<222> (315)..(315)
<223> The 'Xaa' at location 315 stands for Arg, or His.

<220>
<221> misc_feature
<222> (319)..(319)
<223> The 'Xaa' at location 319 stands for Val.

<220>
<221> misc_feature
<222> (329)..(329)
<223> The 'Xaa' at location 329 stands for Thr, or Met.

<220>
<221> misc_feature
<222> (364)..(364)
<223> The 'Xaa' at location 364 stands for Pro.

<220>
<221> misc_feature
<222> (377)..(377)
<223> The 'Xaa' at location 377 stands for Pro, or Ser.

<220>
<221> misc_feature
<222> (380)..(380)
<223> The 'Xaa' at location 380 stands for Ile.

<220>
<221> misc_feature
<222> (403)..(403)
<223> The 'Xaa' at location 403 stands for Arg, or Trp.

<220>
<221> misc_feature
<222> (406)..(406)
<223> The 'Xaa' at location 406 stands for Gly.

<220>
<221> misc_feature
<222> (435)..(435)
<223> The 'Xaa' at location 435 stands for Thr.

<400> 4

Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1 5 10 15

Gln Cys Gly Asn Gln Ile Xaa Ser Lys Phe Trp Glu Val Ile Ser Asp
20 25 30

Glu His Gly Ile Lys Xaa Asp Gly Thr Tyr His Gly Glu Ser Asp Xaa
35 40 45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
50 55 60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65 70 75 80

Xaa Val Arg Xaa Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
85 90 95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
100 105 110

Thr Glu Xaa Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
115 120 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
130 135 140

Leu Gly Gly Gly Thr Gly Ser Xaa Met Gly Xaa Leu Leu Ile Xaa Lys
145 150 155 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Xaa Ser Ser Phe Xaa Val Val
165 170 175

Pro Ser Pro Lys Val Xaa Asp Xaa Val Val Glu Pro Tyr Asn Ala Thr
180 185 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Xaa Cys Ile
195 200 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Xaa Xaa Lys Leu Xaa
210 215 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225 230 235 240

Xaa Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Xaa Ala Asp Xaa
245 250 255

```
Arg Lys Leu Ala Val Asn Met Xaa Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

Ala Leu Thr Val Ala Glu Xaa Xaa Gln Xaa Met Phe Asp Ala Lys Asn
    290                 295                 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Xaa Tyr Leu Thr Xaa Ala
305                 310                 315                 320

Ala Met Phe Arg Gly Arg Met Ser Xaa Arg Glu Val Asp Asp Gln Met
                325                 330                 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Xaa Arg Gly Leu Lys
        355                 360                 365

Met Ala Ala Thr Phe Val Gly Asn Xaa Thr Ala Xaa Gln Glu Leu Phe
    370                 375                 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

Leu His Xaa Tyr Thr Xaa Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

Glu Ala Xaa Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

Glu Ala Tyr Pro Glu Glu
    450
```

<210> 5
<211> 1428
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> CDS
<222> (1)..(1362)
<223>

<400> 5

```
aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga        48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15

cag tgt gga aac caa att ggc tcc aag ttt tgg gaa gtg atc tct gac        96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

gag cac ggc att aag cct gat ggc aca tac cac gga gaa tct gat tta       144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aaa       192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat       240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

tcg gtc cgt tcc ggg cca tac ggg caa ttg ttc cgc cct gat aac tac       288
Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac       336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

act gaa ggc gct gaa ctt gtt gac aat gta cta gat gta gtg cga aaa       384
Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca       432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

ctt gga gga ggt acc gga tcg ggt atg ggc act ctc ctc atc tcc aaa       480
Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145                 150                 155                 160

att cgg gag gag tat cct gat aga atc atg tcc tcg ttc tcc gtt gtt       528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
                165                 170                 175

ccc tca cca aag gtc tcc gac acc gtt gtg gag ccg tac aat gct acc       576
Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att       624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

gac aat gaa gct ctt tat gat att tgc ttc cgc act ttg aaa ctc acg       672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct       720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

ggt gtc act aca tgt ctt cgc ttc cct ggc caa ttg aat gcc gat cta       768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
                245                 250                 255
```

```
cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttt      816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

atg cct ggc ttt gct ccc ctc tct gct aaa ggc gct cag gct tac cgt      864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat      912
Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca      960
Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg     1008
Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg     1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

aac aac gtc aag acc gct gta tgc gac att ccg ccg aga gga ctg aaa     1104
Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

atg gcc gct acc ttc gtt gga aac tca act gcc atc caa gag ctg ttc     1152
Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

aag cgc att tca gaa caa ttc aca gct atg ttc cgc cgc aaa gcg ttc     1200
Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa     1248
Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac cag caa tac cag     1296
Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

gaa gct aca gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa     1344
Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

gag gct tac cct gaa gaa tagacagcag attgtgttgc gttgttcgtt            1392
Glu Ala Tyr Pro Glu Glu
    450

tctctgtgtc aatgcgaaat acacattgat tgcgtt                             1428
```

<210> 6
<211> 454
<212> PRT
<213> Cylicocyclus nassatus

<400> 6

```
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15

Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145                 150                 155                 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
                165                 170                 175

Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
                245                 250                 255
```

```
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

Glu Ala Tyr Pro Glu Glu
    450


<210>  7
<211>  1429
<212>  DNA
<213>  Cylicocyclus nassatus

<220>
<221>  CDS
<222>  (1)..(1362)
<223>

<400>  7
```

```
aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga        48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15

cag tgt gga aac caa att ggt tcc aag ttt tgg gaa gtg atc tct gac        96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

gag cac ggc att aag ccc gat ggc aca tac cac gga gaa tct gac tta       144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aaa       192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat       240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

tcg gtc cgt tcc ggg cca tac ggg caa ttg ttc cgc cct gat aac tac       288
Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac       336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

act gaa ggc gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa       384
Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca       432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

ctt gga gga ggt acc gga tcg agt atg ggc act ctc ctc atc ttc aaa       480
Leu Gly Gly Gly Thr Gly Ser Ser Met Gly Thr Leu Leu Ile Phe Lys
145                 150                 155                 160

att cgg gag gag tat cct gat aga atc ata tcc tcg ttc ttc gtt gtt       528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Ile Ser Ser Phe Phe Val Val
                165                 170                 175

ccc tca cca aag gtc tcc gac acc gtt gtg gag ccg tac aat gct acc       576
Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att       624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

gac aat gaa gct ctt tat gat att tgc ttc cgc act ttg aaa ctc acg       672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct       720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

ggt gtc act aca tgt ctt cgc ttc cct ggc caa ttg agt gcc gat cta       768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Ser Ala Asp Leu
                245                 250                 255
```

```
cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttt      816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

atg cct ggc ttt gct ccc ctc tct gct aaa ggc gct cag gct tac cgt      864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat      912
Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca      960
Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg     1008
Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg     1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

aac aac gtc aag acc gct gta tgc gac att ccg ccg aga gga ctg aaa     1104
Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

atg gcc gct acc ttc gtt gga aac tca act gcc att caa gag ctg ttc     1152
Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

aag cgc att tca gaa caa ttt aca gct atg ttc cgc cgc aaa gcg ttc     1200
Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa     1248
Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac caa caa tac cag     1296
Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

gaa gct acc gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa     1344
Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

gag gct tac cct gag gaa tagaacagca gattgtgttg cgttgttcgt            1392
Glu Ala Tyr Pro Glu Glu
    450

ttctctgtgt caatgcgaaa tacacattga ttgcgtt                            1429
```

<210> 8
<211> 454
<212> PRT
<213> Cylicocyclus nassatus

<400> 8

```
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                   10                  15
```

```
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

Leu Gly Gly Gly Thr Gly Ser Ser Met Gly Thr Leu Leu Ile Phe Lys
145                 150                 155                 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Ile Ser Ser Phe Phe Val Val
                165                 170                 175

Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Ser Ala Asp Leu
                245                 250                 255

Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270
```

```
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

Glu Ala Tyr Pro Glu Glu
    450
```

<210> 9
<211> 1428
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> CDS
<222> (1)..(1362)
<223>

<400> 9

[redacted]

```
aag ttc tct act gca ata atg cgt gag atc gtg cat gta caa gct gga        48
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15

cag tgt gga aac caa att ggt tcc aag ttc tgg gaa gtg atc tct gac        96
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

gag cac ggc att aag ccc gac ggc aca tac cat gga gaa tct gat cta       144
Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

caa tta gaa cga atc aat gtg tac tat aat gaa gca cat gga ggc aag       192
Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

tat gtc ccg cgt gca gtt ctt gtt gat ctc gag ccc gga act atg gat       240
Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

tca gtc cgt tct ggg cca tac ggg caa ttg ttc cgc cct gat aac tac       288
Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

gtg ttt gga cag tct ggc gca gga aat aac tgg gca aaa ggt cac tac       336
Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

act gaa ggc gct gaa ctt gtc gac aat gta cta gat gta gtg cga aaa       384
Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

gaa gct gaa gga tgt gac tgt ctg cag ggc ttc cag cta act cac tca       432
Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

ctt gga gga ggt acc gga tcg ggt atg ggc aca ctc ctc atc tcc aaa       480
Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145                 150                 155                 160

att cgg gag gag tat cct gat aga atc atg tcc tcg ttc tcc gtt gtt       528
Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
                165                 170                 175

ccc tca cca aag gtc ttc gat act gtt gtg gag ccg tac aat gct acc       576
Pro Ser Pro Lys Val Phe Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

cta tcc gtt cat cag ttg gtt gaa aat aca gac gag act ttc tgt att       624
Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

gac aat gaa gct ctt tat gat att tgc ttc cgc acc ttg aaa ctc acg       672
Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

aac cca act tat gga gat ctg aat cat ctt gtg tct gta aca atg tct       720
Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

ggt gtc act aca tgc ctt cgc ttc cct ggc caa ttg aat gcc gat cta       768
Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
                245                 250                 255
```

```
cgt aaa cta gct gtt aac atg gtt cca ttc cct cgt ctt cac ttc ttc      816
Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270

atg cct ggc ttt gct ccc ctc tct gcc aaa ggc gcc cag gct tac cgt      864
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

gct ctt act gta gcc gag cta act caa cag atg ttc gat gcc aaa aat      912
Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

atg atg gcc gct tgc gac cct cga cat gga cgt tat ctc acc gtc gca      960
Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

gcc atg ttc cga gga cga atg agc atg agg gag gta gac gac cag atg     1008
Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

atg tca gtg cag aac aag aac tcc tca tac ttc gta gag tgg att ccg     1056
Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

aac aac gtc aaa acc gct gta tgc gac att ccg ccg aga gga ctg aaa     1104
Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

atg gcc gct acc ttc gtt gga aac tca act gcc att caa gag ctg ttc     1152
Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

aag cgc att tca gaa caa ttc aca gct atg ttc cgc cgc aaa gcg ttc     1200
Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

ttg cat tgg tat act ggt gaa ggt atg gac gag atg gag ttc act gaa     1248
Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

gcc gag tcc aac atg aat gat ctc atc tcc gaa tac cag caa tac cag     1296
Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

gaa gct acc gct gac gat atg ggc gat ctc gat gcg gaa ggc gct gaa     1344
Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

gag gct tac cct gaa gaa tagacagcag attgtgttgc gttgttcgtt            1392
Glu Ala Tyr Pro Glu Glu
    450

tctctgtgtc aatgcgaaat acacattgat tgcgtt                             1428
```

<210> 10
<211> 454
<212> PRT
<213> Cylicocyclus nassatus

<400> 10

```
Lys Phe Ser Thr Ala Ile Met Arg Glu Ile Val His Val Gln Ala Gly
1               5                  10                  15
```

```
Gln Cys Gly Asn Gln Ile Gly Ser Lys Phe Trp Glu Val Ile Ser Asp
            20                  25                  30

Glu His Gly Ile Lys Pro Asp Gly Thr Tyr His Gly Glu Ser Asp Leu
        35                  40                  45

Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala His Gly Gly Lys
    50                  55                  60

Tyr Val Pro Arg Ala Val Leu Val Asp Leu Glu Pro Gly Thr Met Asp
65                  70                  75                  80

Ser Val Arg Ser Gly Pro Tyr Gly Gln Leu Phe Arg Pro Asp Asn Tyr
                85                  90                  95

Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala Lys Gly His Tyr
            100                 105                 110

Thr Glu Gly Ala Glu Leu Val Asp Asn Val Leu Asp Val Val Arg Lys
        115                 120                 125

Glu Ala Glu Gly Cys Asp Cys Leu Gln Gly Phe Gln Leu Thr His Ser
    130                 135                 140

Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys
145                 150                 155                 160

Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Ser Ser Phe Ser Val Val
                165                 170                 175

Pro Ser Pro Lys Val Phe Asp Thr Val Val Glu Pro Tyr Asn Ala Thr
            180                 185                 190

Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Phe Cys Ile
        195                 200                 205

Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
    210                 215                 220

Asn Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Val Thr Met Ser
225                 230                 235                 240

Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu
                245                 250                 255

Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe
            260                 265                 270
```

```
Met Pro Gly Phe Ala Pro Leu Ser Ala Lys Gly Ala Gln Ala Tyr Arg
        275                 280                 285

Ala Leu Thr Val Ala Glu Leu Thr Gln Gln Met Phe Asp Ala Lys Asn
    290                 295                 300

Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala
305                 310                 315                 320

Ala Met Phe Arg Gly Arg Met Ser Met Arg Glu Val Asp Asp Gln Met
                325                 330                 335

Met Ser Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro
            340                 345                 350

Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys
        355                 360                 365

Met Ala Ala Thr Phe Val Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe
    370                 375                 380

Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe
385                 390                 395                 400

Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu
                405                 410                 415

Ala Glu Ser Asn Met Asn Asp Leu Ile Ser Glu Tyr Gln Gln Tyr Gln
            420                 425                 430

Glu Ala Thr Ala Asp Asp Met Gly Asp Leu Asp Ala Glu Gly Ala Glu
        435                 440                 445

Glu Ala Tyr Pro Glu Glu
    450
```

<210> 11
<211> 2655
<212> DNA
<213> Cylicocyclus nassatus

<220>
<221> Intron
<222> (1)..(18)
<223>

<220>
<221> Intron

<222> (76)..(358)
<223>

<220>
<221> Intron
<222> (496)..(637)
<223>

<220>
<221> Intron
<222> (865)..(1374)
<223>

<220>
<221> Intron
<222> (1666)..(1723)
<223>

<220>
<221> Intron
<222> (1915)..(1966)
<223>

<220>
<221> Intron
<222> (2064)..(2119)
<223>

<220>
<221> Intron
<222> (2306)..(2354)
<223>

<220>
<221> Intron
<222> (2475)..(2523)
<223>

<220>
<221> Intron
<222> (2592)..(2655)
<223>

<400> 11
aagttctcta ctgcaataat gcgtgagatc gtgcatgtac aagctggaca rtgtggaaac 60

caaattggtt ccaaggtrcg gtagtttyrr twktytrytg atcgtaattc sggmgytytr 120

dagtrryttt ttycgytgyy ratgttgcat yrtgttgcga taaascyraa aawtcawwag 180

rcgaggctgt aaaagsactt ytacttttra atmcrytgta gcagcatgag tcatcrgcat 240

gtttgcagtg sgtttttttat gcgcwgawcc yctcagaaga tgagaatgcg wtccaytgag 300

cwtagartct grctttctct cgttawctaa ratcaamtta carcrytyca ttttkcagtt 360

ytgggaagtg atctctgacg agcacggcat taagccygay ggcacatacc ayggagaatc 420

tgatytacaa ttagaacgaa tcaatgtgta ctataatgaa gcacatggtt agtcgtayat 480

ccgcttcgtt gtytcccmat gcagrccyct tcagttttta taactgycga aatatcgatc 540

gggctctttt gcagcggccw ygattacgca ataccayygc ygcygcagtg gcrgtcgaaa 600

ttaatgtggt caracgtgaa aatgtggtgc tttyaggagg caartatgtc ccgcgtgcag 660
ttcttgttga tctcgagccc ggaactatgg attcggtccg ytccgggcca tacgggcaat 720
tgttccgccc tgataactac gtgtttggac agtctggcgc aggaaataac tgggcaaaag 780
gtcactacac tgaaggygct gaacttgtcg acaatgtact agatgtagtg cgaaaagaag 840
ctgaaggatg tgactgtctg caggtaaatt tccaagtagt agcaggaaat ggtwtgtgra 900
tagcataaca aaagtcatag aaggaatatg gacgctagtc aaaacaaagw tggacgttar 960
tcggtcgtcc gggacarttt ggaagtcatg ggtcasccaa cacgcttttt tamaagtaca 1020
tcatactctt ttcccacgaa aagctatttt gcgtattacg gggtacaggg gaggggtcaa 1080
aatcacagat tgctgaaaty tggttcactg ragttattgr tgaaaatcat attgattttg 1140
cttgctactg ccttcttttr aggctatgct ttacaatctt ggggcctgga taaccgaatt 1200
gtcygaagtt tttcggtcat cacggacggg gaaggggcat artatcgtta kttcttgkta 1260
tttcgcagca tatggcaatc tytccacttc tgacaagttt tcygtagaaa atatwcttca 1320
aggtstcaag aacyttgctg ctagrgctgt aaaccaayct gtatcycttt cagggcttcc 1380
agctaactca ctcacttgga ggaggtaccg gatcgggtat gggcactctc ctcatctcca 1440
aaattcggga ggagtatcct gatagaatca tgtcctcgtt ctccgttgtt ccctcaccaa 1500
aggtctccga caccgttgtg gagccgtaca atgctaccct atccgttcat cagttggttg 1560
aaaatacaga cgaracttwc tgtattgaca atgaagctct ttatgatatt tgcttccgca 1620
cyytgaaact cacsaaccca acttatggag atctgaatca tcttggtrrg yrayatkcsa 1680
ytgctgagct tdgtrgaatt tvctaattwt ktyhamtdty yagtgtctgt aacaatgtct 1740
ggygtcacta catgycttcg cttccctggc caattgrayg ccgatctwcg taaactagct 1800
gttaacatgg ytccattccc tcgtcttcac ttyttyatgc ctggctttgc tcccctctct 1860
gcyaaaggcg cycaggctta ccgtgctctt actgtagccg agctwacyca rcaggtgcgt 1920
ctgcttatcr ttgwtgayrt gtgtttattc kttgtrtatt ttayagatgt tcgatgccaa 1980
aaatatgatg gccgcttgcg accctcgaca tggacrttat ctcaccgtyg cagccatgtt 2040
ccgaggacga atgagcayga gggtaagtgg mtkmttggyc ytytaryaya rctcrgacga 2100
awtgctgtta tgtcmtagga rgtagacgac cagatgatgt cagtgcagaa caagaactcc 2160
tcatacttcg tagagtggat tccgaacaac gtcaaraccg cygtatgcga cattccgccr 2220
agaggactga aaatggccgc taccttcgtt ggaaacycaa ctgccatcca agagctgtty 2280
aagcgcattt cagaacaatt yacaggttdg tttgtgcaya ttatggtgaa agcagattar 2340
ttgcgaygtt gcagctatgt tccgccgcaa agcgttyttg catyggtaya ctggwgaagg 2400
tatggaygag atggagttca ctgaagccga gtccaacatg aatgatctca tctccgaata 2460
ccarcaatac caggttcggc tgtytttcwt rgayactgtr tttaataatt wttyttgtct 2520

aggaagctac cgctgacgat atgggcgatc tcgatgcgga aggcgctgaa gaggcttacc 2580

ctgargaata gamcagcaga ytgtgttgcg ttgttcgttt ctctrtgtca atgcgaaata 2640

cacattgatt gcgtt 2655

<210> 48
<211> 20
<212> DNA
<213> Synthetic sequence

<220>
<221> misc_feature
<222> (1)..(20)
<223> r is a purine (adenine or guanine); y is a pyrimidine (cytosine, thymine, or uracil)

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is adenine, cytosine, guanine, thymine, or uracil

<400> 48
ggncartgyg gnaaycarat 20

<210> 49
<211> 20
<212> DNA
<213> Synthetic sequence

<220>
<221> misc_feature
<222> (1)..(20)
<223> r is a purine (adenine or guanine); y is a pyrimidine (cytosine, thymine, or uracil)

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is adenine, cytosine, guanine, thymine, or uracil

<400> 49
gaytgyytnc arggnttyca 20

<210> 50
<211> 20
<212> DNA
<213> Synthetic sequence

<220>
<221> misc_feature

<222> (1)..(20)
<223> r is a purine (adenine or guanine); y is a pyrimidine (cytosine, thymine, or uracil)

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is adenine, cytosine, guanine, thymine, or uracil

<400> 50 20
tgraanccyt gnarrcartc